Title: Evaluating university teaching and learning in an outcome-based model: replanting Bloom

Author: Maureen Mary Morris

Year: 2008

Repository DOI:

Evaluating university teaching and learning in an outcome-based model: replanting Bloom

Maureen Mary Morris
University of Wollongong

Evaluating University Teaching and Learning in an Outcome-Based Model: Replanting Bloom

A thesis submitted in fulfilment of the
requirements for the award of the degree

DOCTOR OF PHILOSOPHY

from the

UNIVERSITY OF WOLLONGONG

By

Maureen Mary Morris
Bachelor of Science/Diploma of Education, University of Western Sydney
Master of Arts (Pure Mathematics), Sydney University

School of Mathematics and Applied Statistics
2008

Declaration

In accordance with the regulations of the University of Wollongong, I, Maureen Mary Morris, declare that this thesis, submitted in fulfilment of the requirements for the award of Doctor of Philosophy in the School of Mathematics and Applied Statistics, University of Wollongong, is my own original work, except where due references are made. It has not been submitted for a degree at any other university.

Maureen Mary Morris

Abstract

This thesis was inspired by an experienced teacher's desire to enhance student learning through implementation of a teaching/learning framework focused on promotion of higher order cognition. Two case studies document the construction, implementation and evaluation of learning frameworks for two disparate undergraduate university subjects.

Structurally, the thesis falls into three component parts. In the first, the researcher has reviewed the literature for an appropriate methodology, grounded her understanding of student learning through examination of relevant learning theories, canvassed suitable pedagogical strategies before construction of the teaching/learning frameworks, and devised an evaluation framework. In the second part, the two case studies have been described and in the final part, the threads of evidence have been drawn into the conclusion.

Action research afforded an appropriate methodology for the study. It offered facility for a spiral of implementation, review and re-implementation. Bound as a practitioner by the pragmatic perspective of *what works*, the researcher engaged multiple methodologies (grounded research encompassing elements of phenomenology and ethnography) in both case studies. She adopted a mixed method approach, with evidence derived from assessment data, survey responses, her annotated journal and comments from collaborating teachers and students.

The researcher's primary intent was to construct *aligned* teaching/learning frameworks that promoted contextualised thinking for students in the two disciplines. Judgment of the effectiveness of the resulting frameworks in enhancing student learning required a strict evaluative regimen.

Key issues percolated through the thinking of the researcher/teacher:

- *life-long learning;*
- *meta-cognition* and *deeper learning;* and

- marking of assessment that recognises achievement of learning objectives, offers students task related feedback and does not merely represent an aggregation of marks for ranking of students along a curve.

Therefore, strategies were included that fostered independent learning and promoted productive collaboration, while marking criteria formed the focus for aligning marking with the objectives.

The primary case study examined teaching and learning in a foundation course in statistics at the University of Wollongong in Australia. The intent was to foster *statistical thinking* in students. Experienced in the field, the teacher assumed an active role as a participant researcher. In consultation with discipline experts and innovative teachers, the researcher/teacher observed the existing environment for a single session (N=159). Learning objectives were then rigorously scrutinised and behaviourally reframed; objectives were specified for learning and assessment tasks; and marking criteria devised to scaffold student responses, check assessment for objective achievement and provide detailed and task related feedback. Thus the objectives formed the agents of *constructive alignment*.

Implementation of the selected strategies was tracked over the subsequent four sessions (cohorts ranging in size from 152 to 192 students). Evidence of student learning and the effectiveness of the framework was derived from:

- assessment marks and grades;
- deconstruction of assessment tasks and responses using the revised Bloom's Taxonomy (Anderson and Krathwohl, 2001);
- student survey responses;
- teacher and marker survey responses;
- the researcher's journal, annotated by collaborative teachers; and
- peer discussions.

Results have highlighted increases in mean marks in summative assessment accompanied by shifts to higher order cognitive demand in assessment tasks across the implementations. Furthermore, strong correlations between proportions of students reporting confidence in topic learning and exam performance have lent credence to the teacher's claim that *students know what they know* and *know what they do not know.*

The aim of the second case study was the design and implementation of an aligned curriculum for a subject focused on promoting *critical and evaluative thinking* in undergraduate accounting students. Although not the teacher/researcher's field of expertise, intense consultations with the subject designers produced behaviourally framed objectives and a teaching/learning framework that targeted the desired skills. This case study consisted of a single implementation (N=223). Results were not conclusive, but examination of the detail has provided fresh insight into the potential value of peer evaluation and student portfolios to address the desired thinking.

Comparison of the two case studies has highlighted the marked similarity between the teacher's expectations of *statistical thinking*, which underpins the University of Wollongong subject, and *critical and evaluative thinking*, which underpins the University of Western Sydney subject. 'Structure' has been identified as essential to successful implementation of the frameworks targeting discipline *thinking.* The structure of the desired thinking needs not only to be modelled but also to be recognised by students before it is effectively assimilated.

The researcher's journey has required reflective practice that includes both telescopic and microscopic review of her thinking, her habits and the action and reaction occurring within her classroom. The evaluation of student learning undertaken in this thesis has formalised the teacher's informal and intuitive response to the ostensibly absurd behaviours that take place as her students learn. Her deconstruction and interpretation of the apparent incongruities has at once affirmed past practice and inspired its renewal.

Acknowledgements

Although the onus of research is predicated on a journey of solitary thought, and the researcher is proud of her accomplishment, she acknowledges that this thesis has emerged in an environment redolent with support. She wishes to thank her supervisors for their encouragement, ideas and criticisms. Professor David Griffiths’ humour, intellect and expertise with a red pen have spurred this writer to aim high and to treat the English language with the respect it richly deserves. Dr Anne Porter is an inspired educator who has fired the researcher with her enthusiasm for teaching statistics and treated this teacher as a valued colleague.

Born of parents who perceived education and independent thought to be more precious than wealth, brother, Denis and sister, Colleen have fuelled the drive to gain the competitive edge in academic achievement. Their encouragement, founded in expectation and belief in my ability, has been immeasurable.

I thank my children who have come to view Bloom as an annoying cousin who came to dinner and refused to leave. A gifted writer, Patrick has served many hours in review of my ideas and now regards himself expert on Bloom’s taxonomy. Peter’s training as a teacher has both been supported by his secondment as an evaluator of my theoretical perspectives and repulsed by the volume of those ideas. My daughters Louise and Kate have shown great restraint in turning a blind eye to their mother’s housekeeping and have encouraged their eccentric mother in her academic pursuits. Louise’s assistance with typing is appreciated as the researcher acknowledges her ‘one finger’ entry of this entire thesis has been daunting.

I thank my husband who has long since lost touch with my ideas and become accustomed to an absent minded wife and the invading dust mites that occupy the vast piles of rejected copy.

I acknowledge Anne Harper, Melissa Bennett, David Guy and Joe Tiziano who have provided the technological expertise that has saved this thesis from destruction on numerous occasions. I thank my friends Carole, Elahe, Raed and Rebecca (among others) who have suffered the same struggle to produce a thesis.

Last, but not least, I acknowledge the profound support provided by Emmie, Toby, Moghul and Georgie, my spaniels. They have suffered every tragedy and shared every joy in this epic journey. They have blindly accepted my idiosyncrasies and treasured my time with them as though it was for their benefit alone. Their companionship and their provision of the warmth of their bodies have saved me from cold and loneliness in the long periods of solitude.

Table of Contents

Chapter 4

Teaching for learning: as easy as falling off a horse!

Chapter 5

Evaluation

Learning from teaching: Looking for atoms of meaning

Chapter 6

Case Study

Doing the sums- improving a student's perspective of statistics

Tables

Figures

Chapter 1
Introduction
Informing practice: in pursuit of the elusive rabbit

… suddenly a White Rabbit with pink eyes ran close by her.
There was nothing so *very* remarkable in that; nor did Alice think it so *very* much out of the way to hear the Rabbit say to itself, “Oh dear! Oh dear! I shall be too late!” (when she thought it over afterwards, it occurred to her that she ought to have wondered at this, but at the time it all seemed quite natural); but when the Rabbit actually *took a watch out of its waistcoat-pocket*, and looked at it, and then hurried on, Alice started to her feet, for it flashed across her mind that she had never before seen a rabbit with either a waistcoat-pocket, or a watch to take out of it, and burning with curiosity, she ran across the field after it, and was just in time to see it pop down a large rabbit-hole under the hedge.

In another moment down went Alice after it, never once considering how in the world she was to get out again.

The rabbit-hole went straight on like a tunnel for some way, and then dipped suddenly down, so suddenly that Alice had not a moment to think about stopping herself before she found herself falling down what seemed to be a very deep well.

Either the well was very deep, or she fell very slowly, for she had plenty of time as she went down to look about her, and to wonder what was going to happen next. First, she tried to look down and make out what she was coming to, but it was too dark to see anything; then she looked at the sides of the well, and noticed that they were filled with cupboards and bookshelves: here and there she saw maps and pictures hung upon pegs.

… Down, down, down. Would the fall *never* come to an end? “I wonder how many miles I have fallen this time?” she said aloud. “I must be getting somewhere near the centre of the earth. Let me see: that would be four thousand miles down. I think –” (for, you see, Alice had learnt several things of this sort in her lessons in the schoolroom, and though this was not a very good opportunity for showing off her knowledge, as there was no one to listen to her, still it was good practice to say it over) “– yes, that’s about the right distance – but then I wonder what Latitude or Longitude I’ve got to?” (Alice had no idea what Latitude was or Longitude either, but they were nice grand words to say.)

… Down, down, down. … when suddenly, thump! thump! down she came upon a heap of sticks and dry leaves, and the fall was over.

Alice was not a bit hurt … and the White Rabbit was still in sight … There was not a moment to be lost; away went Alice like the wind, …

From "Down the Rabbit-Hole" in
Alice's Adventures in Wonderland by Lewis Carroll, pp. 3-5.

1.0 Inspiration: teaching in Wonderland

The researcher has enjoyed many years as a classroom teacher. This enjoyment has been fuelled by challenges arising from the eclectic blend of personalities, perspectives, interests and goals that constantly bombard the orchestrator of a learning environment. As a teacher, the researcher has approached each lesson armed with the determination to promote fulfilment of the predetermined learning objectives. Despite her focused preparation, student responses have frequently necessitated some form of *deconstructive translation* in order to further communication of necessary concepts. At times it has appeared to the researcher that her classrooms have assumed the dimensions of the Mad Hatter's Tea Party (Carroll, 2002). As a researcher, the teacher must, like Alice, use her powers of observation and wit to record the *absurdities* encountered in her journey as she attempts to unravel their import in enhancing student learning.

This study chronicles the journey of the teacher as she develops the reflective and evaluative perspective of her profession as a researcher. Throughout its writing the dual nature of the writer is evidenced. The researcher has emerged through considered study of the literature and systematic evaluation of the teacher's practice while the teacher has been revealed in the perspective of experienced practitioner through her observations and reflective grounding of the literature in her experience. There were times, however when the perspectives merged and the image of the teacher and the researcher became indistinguishable.

1.0.1 Active teaching: a watching brief

Although students appear to construct meaning framed by their own backgrounds, abilities and interests, the learning framework in which this learning takes place has been contrived by the teacher. It is the contention of the researcher that an effective design, founded on well defined expectations of student learning, has the facility to guide their collective learning to achieve common objectives.

An initially reluctant participant in the introduction of the outcome-based curriculum implemented in New South Wales high schools in the 1990's, the teacher came to value its merits. Not only did it appear to promote accountability (an important notion for a student conception of fairness), but also fostered *meta-learning* through its mandated specification of the outcomes targeted by assessment and reporting of achievement against those outcomes. The teacher was also working at university and observed repeatedly that her university students, products of the high school curriculum, were confounded by examinations and assignments that seemed poorly connected to the discipline material covered in lectures and tutorials. The question arose: Could these issues be addressed by an outcome-based model?

This study arose more particularly from the teacher's classroom experience both in high school mathematics classes and in university mathematics and statistics classrooms. Her professional practice has afforded much opportunity for observing students as they learn. Her reflective ruminations of what she has done and what her students have achieved (or indeed, not achieved) have constantly fuelled fresh approaches.

As with many other dedicated teachers of mathematics, this teacher has laboured under the perception that while her efforts to get her students *doing* mathematics had met with reasonable success, many could not *carry their knowledge* to the next year (nor indeed into the next assessment task!). These students demonstrated the ability to *do* the problems/exercises at or near the time of instruction and practice, but they did not *know* the underlying concepts sufficiently to retain and transfer knowledge and skills in the longer term. Their learning was *shallow*. This was also evident in the university classrooms. There, many students in first courses in mathematics or statistics demonstrated poor basic skills, even though their high school records had been indicative of high achievement!

Working with weak students in high school mathematics has also led the teacher to believe that the less able students were at grasping fundamental concepts, the more adept they were at remembering *nonsense* processes and skills. Even for more able students in an 'examination focused' learning environment, memorising poorly understood processes appears to offer the easiest path to 'passing' and hence drives their *shallow* learning. Promotion of this type of mathematical learning is a scary prospect as it is so limiting! Such learning appears to arrest student motivation and engagement. It is also prohibitive of extending their mathematical development, as fundamental understandings are essential building blocks for more complex concepts and skills.

Although primarily a teacher of mathematics, the teacher also has had extensive experience in high school science classrooms. Her experience there has affirmed her belief that an experiential approach to learning that employs enquiry-based tasks effectively promotes acquisition of the necessary concepts and skills. These styles of learning have much to offer in the mathematics (and particularly, the statistics) classroom.

Marriage of discipline knowledge and skills with practical (and relevant) application is appropriate in fostering the transfer of that knowledge and those skills. If the aim of the good teacher is to promote engagement of the higher cognitive skills demanded by a *deep approach* to learning, then students need to be more actively engaged in their own learning (Biggs, 1999). Reflections on her own learning, her teaching experience and the literature persuaded the researcher that these notions should underpin an appropriately designed pedagogy.

In many university classrooms, the teacher's transfer of discipline-based knowledge has been the paramount consideration and has evidenced little recognition of knowledge of student cognitive development (Yorke, 2001).

> Experience suggests that most programs in higher education are based on a set of general assumptions in which the subject discipline, rather than student development, is dominant. (Yorke, 2001, p. 118)

The rapid development of technology and acceleration in the growth of human knowledge has seen a burgeoning need for students to become *life-long* learners. (West, 1997) Whilst the teacher recognises the importance of discipline knowledge and skills, she has also come to realise that if these alone form the focus for intended student learning, the result can be a 'blinkered' perspective of their use.

There is a need to augment these skills with the higher order cognitive skills that facilitate their *synthesis*, *transfer*, and extended *application*. Students need to develop the capacity to use them to *critically evaluate*, *form judgements* and *creatively design solutions to real life problems* in emulation of *expert thinking* (Te Wiata, 1996). They need to take greater responsibility for their own learning, developing the capacity to organise and critically reflect on their own learning. (Toohey, 1996) Only this type of *meta-learning* can equip them as *life-long* learners. Addressing this type of learning requires Biggs' notion of *good teaching*.

> … a view of teaching that is not just about facts, concepts and principles to be covered and understood, but about:
>
> 1. What it means to *understand* those concepts and principles in the way we want them to be understood.
> 2. What kind of TLAs (teaching/learning activities) are required to reach those kinds of understandings. (Biggs, 1999, p. 63)

Harbouring a long term commitment to reflective practice, and goaded by the lofty ambitions of being a *good teacher*, the teacher needed to move beyond the informal observations that had previously afforded little more than anecdotal evidence to propel her perceptions of effective teaching. The teacher needed to become the researcher to provide appropriate evidence in her quest to improve her students' learning and to systematically evaluate the classroom action.

1.0.2 Undertaking the quest

Armed with a dual perspective, the teacher/researcher embarked upon this journey with the same spirit of curiosity and adventure as did Alice (Carroll, 2002). *Her* quarry has been the learning of her students and it has proved equally as elusive as Alice's White Rabbit. Ultimately, as with Alice's journey, greater *lucidity* is restored in the *dénouement*. The final deconstruction and analysis of the apparently *absurd* and *asynchronous* experiences gained throughout this journey has been to shed some light on the students' learning experiences in the teacher's classroom.

1.1 Intentions

The concept of learning is complex. Reflective practice supports teacher understanding of the processes involved in the learning of students. But valid reflection should be grounded in the collective wisdom and experience afforded by reference to the literature. In planning this study, the researcher's deliberations identified the following aspects as fundamental to a research study of a functioning learning environment:

- methodological considerations (discussed in Chapter 2);
- an understanding of the concept of learning (explored in Chapter 3);
- a perspective of current pedagogical design (provided in Chapter 4) and
- an appropriate framework for evaluation (devised in Chapter 5).

The research presented in these chapters was completed after review of both current and historical perspectives. This process did not result in radical changes to the teacher's values and beliefs. It did however broaden and deepen her understanding. It offered new vistas and revealed new possibilities for practice.

The teacher believes that alignment of teaching, learning and assessment will enhance student learning. Alignment can be achieved through clear definition of the expected learning. Since assessment provides a clear focus of how teachers expect students to demonstrate their learning, it affords instrumentation for the alignment. Through transparent specification of how expected learning might be demonstrated and hence evaluated, all facets of teaching and learning in the classroom context can be focused.

It is not sufficient for the learning objectives to be implicit. Teachers devise assessment that they believe tests students against the subject learning objectives, but marking of the assessment should also align with the targeted knowledge and skills, and students should *know* what is expected of them, not only through overt specification of the task objectives, but also through reference to the experience and modelled thinking to which they have been exposed in the classroom. Thus alignment is only achieved through the complex orchestration of all facets of the teaching, learning and assessment experiences. (delMas, 2002; Biggs, 1999)

1.1.1 Research design: addressing the '*ologies*:

Three questions underpin the research design of this project:

1. (The ontology) How does the researcher view the nature of the reality of the classroom?
2. (The epistemology) How does she view the nature of knowledge and is there some connection between the researcher and this knowledge?
3. (The methodology) How does she seek the truth of classroom reality? (Mertens, 2005)

Initially these questions appeared not to have clearly defined answers. However maintaining a focus on what the teacher looks for in the classroom, an answer appeared to emerge. Most teachers work with the reality before them. Common sense and practicality define classroom success. A teacher's evaluation of effective practice is founded on success as reflected in student achievement of the predetermined learning objectives, that is, the teacher's guiding set of concepts are pragmatic. In this study, it is the researcher's hand that has framed the evaluation in terms of the values she has acquired through professional practice and reading, and her evidence has arisen from both qualitative (annotated commentary and discussion) and quantitative (survey and assessment detail) sources.

1.1.2 Defining learning

Although a trained teacher, the researcher's recollections of learning theory and pedagogical design had lost specificity and had blended with experience over her years of practice. This study provoked an intensive audit and review of the relevant literature. The process has heightened recollection of previously encountered theory and illuminated professional practice obviously inspired by it. One such example figured prominently. The researcher has become accustomed to deconstructing her own mathematical thinking in order to *scaffold* such thinking for her students. Revisiting the work of Ausubel provided the historical basis! Inspired by his work (Ausubel, 1960; Ausubel, 1968; Ausubel, 1978), the researcher has woven his notions of *scaffolding* and *organisers* into the design her teaching/learning framework for this study.

Formulation of specific learning objectives and construction of the teaching/learning framework can only been achieved through appropriate classification of the desired knowledge and skills. The description of student learning has thus necessitated the examination of appropriate taxonomies.

1.1.3 A blueprint for learning

Biggs (1999) claimed that

> *Good teaching is getting most students to use the higher order cognitive level of processes that the more academic students use spontaneously.* (Biggs, 1999, p. 58)

and advocated a pedagogy that

> … maximises the chances of engaging students' learning processes in this way. (Biggs, 1999, p. 58)

The researcher had used both expository and experiential forms of teaching in the classroom. Her experience had identified the latter as more likely to produce a *deeper learning*. Her review of the literature and consequent exposure to strategies less familiar to her (particularly in the University of Western Sydney case study discussed in Chapter Seven) have given rise to innovative approaches framed to promote this type of learning.

Fundamental to an effective teaching/learning framework is alignment of teaching, learning and assessment. Clear specification of the learning objectives can afford a mechanism for ensuring that the nominated knowledge and skills are targeted by the instructional strategies and are the ones assessed in determining achievement of the objectives. Such alignment is facilitated through application of appropriate taxonomy. In this study, a revision of the taxonomy of Bloom (Anderson, and Krathwohl, 2001) has been used to this end.

Students need to understand what they are expected to learn and hence learning objectives need to be clearly formulated so that assessment not only addresses the learning teachers expect but also the learning that students believe they need to demonstrate. (Biggs, 1999) Effective assessment then can provide a means for focusing the teaching and learning.

In designing a learning framework supportive of achievement of the desired learning and preparing students for assessment, the researcher targeted the following:

1. The subject presentation incorporating:
    - Experiential learning;
    - Defined objectives;
    - Collaborative learning;
    - Authentic tasks;
    - Use of technology;
    - Modelled thinking;
    - Portfolios/workbooks;
    - Marking Guides; and
    - Formative assessment.
2. Construction of an aligned pedagogy:
    - Taxonomy
    - Behavioural objectives.
3. Structured evaluation of the framework.

1.1.4 Evaluation: a blueprint for detecting learning

In her practice the teacher has endeavoured to design assessment that checks achievement of her learning objectives. To ensure that this detection is not too late to rectify misconceptions and address shortfalls in understanding, she has been accustomed to scanning student behaviours for evidence of their learning throughout the teaching/learning process. Much of a teacher's design and evaluative processes are internalised and hence informal. In this study, the researcher has needed to systematically organise the pedagogical design and evidence collecting for more formal analysis. To this end, the researcher embarked upon the search for an evaluative framework that would provide structure and keep the entire process *honest*. Thus once the learning objectives had been specified, such a framework would maintain focus on

- the design of a teaching learning framework that promotes achievement of these objectives;
- the design of assessment that checks achievement against these objectives and
- the identification, collection and analysis of evidence of what *worked* and what *did not work.*

Bain (1999) has recommended that scholarly educational studies should

- address interventions of educational significance that afford new perspectives;
- be grounded in educational and evaluation literatures;
- scrutinise both the learning outcomes and the path to their achievement and
- draw conclusions from firm evidence.

The researcher has drawn upon these notions to focus and guide the teacher's perception of the value of her teaching/learning frameworks to her students learning.

1.2 Rationalisation

The primary aim of the teacher throughout this study has been the promotion of higher order thinking in her students. She has sought to develop students beyond the acquisition of the fundamental knowledge and skills of their disciplines to enable them to use ‘expert’ thinking to solve real world problems beyond the context of the classroom. In order to ensure that her judgement of effective teaching to this end is evidence-based, the researcher has trawled the wealth of literature in support of the teacher’s design, and to establish an evaluative framework that will facilitate the capture and analysis of appropriate evidence. These aspects of the study have reinforced the teacher’s informal practice and grounded it to enable extension of the processes involved beyond her own classroom.

1.2.1 An innovation or a renovation?

The foundations for this study lie in the experiences of the teacher. From the outset of her training as a high school teacher, her lessons have been planned around behaviourally framed objectives. Her practice has been strongly influenced by her study of Bloom and his colleagues (1974) in her teacher training. Her university teaching experiences have supported Biggs’ contention that

> Constructive alignment is common sense, yet most university teaching is not aligned. This is possibly because many academics, holding traditional transmission theories of teaching that ignore alignment, simply haven’t seen the need to question their assumptions. (Biggs, 1999, p. 73)

The teacher’s understanding of the merit of defined learning outcomes and her experiences in implementing an outcome-based curriculum in high schools has inspired an attempt to design an aligned pedagogy for implementation in universities. In so doing, she has realised that this study may not necessarily prove *innovative*, but rather a *renovation* of what *has worked* effectively in many classrooms in the past. However in presenting an aligned teaching/learning framework and forming judgement of its effectiveness in the university environment, she has demonstrated some measure of innovation in her evaluative framework, in particular with respect to her evaluation of student learning from multiple perspectives.

1.2.2 A job worth doing?

delMas (2002), referring to reform in statistics education, has acclaimed current research for its strong grounding in the literature and expertise of the researchers. He has endorsed its potential to improve statistics learning through implementation of the researchers' recommendations. He has however, called for studies that seek to evaluate effectiveness of the research claims.

This study has attempted to redress this deficit this in some measure. It provides detailed evaluations of student learning in aligned pedagogies in two diverse disciplines: foundation studies in statistics and critical and evaluative appreciation of accounting models. Despite the obvious differences in discipline content, some remarkable similarities in the thinking required of students emerged as the study progressed. Evidence gathering has been a key focus for both studies.

1.2.3 At journey's end?

The teacher has taught for many years. From its beginnings, the journey of the researcher has drawn upon the teacher's capacity to critically reflect upon her own practice in terms of its impact upon her students' learning. Throughout the journey, the researcher has endeavoured to more systematically inform her professional practice as a teacher. Her growth is evident in this record of the journey. The growth has resulted from exposure to her own achievements (and perhaps as frequently, her failures) in the classroom, and also to an eclectic mix of ideas and actions from the influential 'others' who have shared her journey, in person or through their writing. However her growing is also far from over. New students, new teachers and new ideas still bombard her practice, but this journey as a researcher has served to hone her ability to more effectively evaluate that practice.

1.3 Framing the purpose: aims and objectives

This study has sought to inform professional practice in higher education through application of an outcome-based model of teaching and learning. It has aimed to investigate the contention that a pedagogy aligned through specified objectives enhances student learning in the selected undergraduate university subjects. Evidence in investigating this contention was sought from two distinct case studies:

1. Exploring Variation and Uncertainty in Data (known throughout this study as STAT131): This study tracked implementation of the learning framework devised for a foundation statistics subject across five successive sessions at the University of Wollongong;
2. Introduction to Accounting Theories and Philosophies (200102) (known throughout this study as: Introduction to Accounting Theories and Philosophies): This study reports on implementation of a learning framework devised to promote critical and evaluative thinking in final year accounting students at the University of Western Sydney.

The researcher subsequently examined the results from the specific contexts of the case studies and sought commonality for more generalised support of the aim.

In examination of the case studies, the researcher focused on each of the sets of more subordinate claims specified in Table 1.1.

Table 1.1: Objectives of the case studies incorporated in this thesis

Objectives	Case Studies
Clear specification of the learning objectives focuses *teaching* on student achievement of the desired levels of knowledge and skills; Clear specification of the learning objectives focuses *student learning* on achievement of the desired levels of knowledge and skills; Clear specification of the learning objectives focuses *assessment* on student achievement of the desired levels of knowledge and skills; and Clear specification of learning objectives promotes alignment of *teaching, learning* and *assessment*.	STAT131 and Accounting Theories and Philosophies

Table 1.1 continued: Objectives of the case studies incorporated in this thesis

Objectives	Case Studies
Alignment of teaching, learning and assessment enhances the development of *statistical thinking* in a fundamental statistics subject for university students.	STAT131
Alignment of teaching, learning and assessment enhances the development of *critical and evaluative thinking* in undergraduate accounting students at university.	Accounting Theories and Philosophies

1.4 Organising the narrative

This chronicle does not represent a sequential reporting of events in the study. It has rather followed a logical definition of terms and review of relevant literature (Chapters 2 to Chapter 5), a presentation of the primary case study (Chapter 6) and a presentation of a further case study in another discipline and ostensibly examining different learning outcomes (Chapter 7).

The growth of the researcher/teacher is evidenced in the development of the case studies. Chapter 7 was written before the completion of Chapter 5 (in which the development of the evaluation framework is discussed) and although the evidence-based approach has been implemented, the more complex structure of Chapter 6 required refinement of the evaluation framework before its completion.

The researcher's review of the literature has broadened the teacher's perspective, highlighting fresh approaches to instructional activities targeting *meta-cognition*. The researcher's forensic examination of the evidence has encompassed sources previously not considered by the teacher, and has detected a burgeoning image of students' abilities to judge the extent of their own learning.

Chapter 2
Methodology: A case to answer?

"Now, Kitty, let's consider who it was that dreamed it all. This is a serious question, my dear, and you should *not* go on licking your paw like that – as if Dinah hadn't washed you this morning! You see, Kitty, it *must* have been either me or the Red King. He was part of my dream, of course – but then I was part of his dream, too! *Was* it the Red King, Kitty? You were his wife, my dear, so you ought to know – Oh, Kitty, *do* help settle it! I'm sure your paw can wait!" But the provoking kitten only began on the other paw and pretended it hadn't heard the question.

Which do *you* think it was?

From "Which Dreamed It?" in
Through the Looking Glass by Lewis Carroll, pp. 237-238.

2.0 Qualitative research: a view from inside the rat's maze

> Qualitative research is multimethod in focus, involving an interpretive, naturalistic approach to its subject matter. This means that qualitative researchers study things in their natural settings, attempting to make sense of, or interpret, phenomena in terms of the meanings people bring to them. Qualitative research involves the studied use and collection of a variety of empirical materials – case study, personal experience, introspective, life story, interview, observational, historical, interactional, and visual texts – that describe routine and problematic moments and meanings in individuals' lives. (Denzin and Lincoln, 1994, p. 2)

The study of problems in context may impose constraints upon the researcher which are quite different in nature to those imposed on the experimental researcher. Not only does the researcher lack control over impacting variables in a complex and interactive environment, but determining solutions to problems arising in such contexts may even require their pervasive presence. Collecting evidence will incorporate multiple perspectives and sources in an attempt to objectify the reality, and will necessitate declaration of the position of the researcher as it is their mind's eye view that records and interprets the action. Hence the fruits of the researcher's efforts represent a

> … *bricolage*, a complex, dense, reflexive, collage like creation that represents the researcher's images, understandings, and interpretations of the world or phenomenon under analysis. (Denzin and Lincoln, 1994, p. 2)

The qualitative researcher's methodological selections are constrained by the context.

> The choice of which tools to use, which research practices to employ is not set in advance. The (choices will depend on) … what is available in context, and what the researcher can do in the setting. (Denzin and Lincoln, 1994, p. 2)

In this study, the researcher has been actively situated within a dynamic teaching/ learning context. Hence the practice of an experienced teacher has been stamped upon this project. The researcher's role however, has extended beyond a teacher's guiding hand in a social constructivist learning environment. Her experience, beliefs and practices have underpinned not only the teaching activities but also inform the evaluation strategies.

2.0.1 Behind the teaching: A teacher's perspective of knowledge

Mertens (2005) has described the questions that give rise to the specification of the research paradigm in the following way:

1. What is the nature of reality? - Ontology
2. What is the nature of knowledge and the relationship between the knower and the would-be-known? - Epistemology
3. How can the knowledge and understanding be acquired? – Methodology.

The practice of teaching is predicated upon the answers to these questions, since the prime business of the teacher is designing curricula that will facilitate the acquisition of knowledge and understanding by students. The aims and objectives of this project are commensurate with those of 'good teaching'. Hence it is pertinent here to examine some of the recognised aspects of 'good teaching' and their impact on the design of this research.

Entwistle et al. (2000) explored conceptions of 'good teaching'. They examined teachers' own conceptions of 'good teaching' and concluded that these have arisen out of each individual's experiences and history. They recognised that

> … a fundamental change in thinking about the nature of knowledge led to an equivalent change in ways of teaching. (Entwistle et al., 2000, p. 7)

and have provided a diagrammatic representation linking epistemological changes to developmental trends in thinking and conceptions of teaching (Figure 2.1). The two most readily identifiable approaches to teaching delineate the developmental continuum. A teacher-focused approach, centred upon discipline content, envisages an absolute reality that is transmitted from the expert to the novice through exposition. On the other hand, a student-centred approach, focused on interactive engagement, recognises multiple realities that are socially constructed and value laden. (Entwistle et al., 2000)

Figure 2.1 depicts this developmental transition of increasing awareness and conceptual complexity alongside a teaching perspective continuum ranging from more traditional approaches to university teaching by discipline experts to current trends aimed at fostering the independent and lifelong learning envisaged by regulatory bodies providing the policies and finance (West, 1998).

Please see print copy for Figure 2.1

Figure 2.1: Diagrammatic representation of developmental trends in thinking and conceptions of teaching
(Entwistle et al., 2000, p. 7)

Entwistle et al. (2000) concluded that simpler conceptions of 'good teaching' reflect an

> … unrealistic over-simplification of experience. (Entwistle et al., 2000, p. 22)

whereas

> Sophisticated conceptions of teaching … (are) always derived from a thoughtful consideration of past experience. Such conceptions imply both an expanded awareness of the relationship between teaching and learning and the strategic alertness to classroom events. (Entwistle et al., 2000, p. 23)

They believed that such 'sophisticated' perceptions of 'good teaching' arise from a teacher's belief in an empathetic partnership with the learner. This capacity to empathise enables the teacher to reflect on an integrated pattern of experiences (both personal and those of the students), of the discipline content and how best to 'teach' it, and of classroom interactions. Such deliberations facilitate the teacher's focus on the learners' optimal path to achievement of a devised learning goal. (Entwistle et al., 2000)

The researcher has had long experience to draw upon, both in high school and university mathematics and statistics classrooms. Her conceptions of 'good teaching' have been experientially *informed* by her action, observations, reactions and reflections and have been evidenced in her interactive engagement with her students in their learning. Her social constructivist beliefs define her practice and enable the learning of her students.

2.0.2 Selecting the research paradigm

Conscious then of the researcher's adherence to social constructivist learning theory and the need to address the aims and objectives specified in Chapter 1, identification of a paradigm for this study has involved examination of the several options compared in Table 2.1. At the onset of this study, the researcher's position was pragmatic, and the methodology virtually determined by the goals espoused (Chapter 1) and the evidentiary sources available to her. However, as a consequence of the researcher's perception of knowledge as socially constructed, the possibility of re-examining this position has also been accommodated with illumination provided by reflective evaluation as the study has evolved.

2.0.3 Fit to purpose: the aims and objectives

As specified in Chapter One, paragraph 1.3, this study 'has aimed to investigate the contention that a pedagogy aligned through specified objectives enhances student learning in the selected undergraduate university subjects'. In particular, through overt specification of the learning objectives, it has sought to construct a teaching/learning framework that would:

1. align teaching, learning and assessment;
2. focus teaching, learning and assessment towards achievement of the objectives; and
3. improve statistical learning (Chapter 6) and/or critical and evaluative thinking (Chapter 7).

Fundamental to achieving the aims of the study has been selection of a research methodology that has engaged appropriate evidence gathering in the complex social dynamics of a classroom learning environment. There have been multiple perspectives of the reality of learning in those classrooms:

- the learning observed by the teacher;
- the learning perceived by the individual students;
- the learning demonstrated by students in the assessment; and
- recognisable shifts in the cognitive demand of assessment subsequent to the changes in pedagogy.

These multiple perspectives have given rise to multiple sources of evidence and these have included observations, written commentary, surveys, assessment data and deconstruction of assessment in order to classify cognitive demand.

In practice, the teacher has selected teaching/learning strategies that have *worked*, and eliminated or revamped those that have not. Her arbitrations have been made in the context of *her* students' learning needs. Hence for the researcher, her 'reality' has emerged within the classroom context and because of her role as a participant teacher, that 'reality' has also been impacted by her own values and beliefs. *What works* in terms of student learning within her classroom has emerged from the teacher's experience and history and the researcher's education and reading. Hence her methodological approach to this study has consequently been pragmatic. However although expedience has driven the pedagogical designs peculiar to her selected classrooms, the inclusion of two disparate case studies has been made to support extrapolation beyond those studies and hence with the intention of generating more generalised conclusions.

Table 2.1: Research paradigms

Paradigm	Ontology	Epistemology	Methodology	Notes
Positivist	One knowable, value-free reality	Observable Perceived independence of researcher and researched Researcher controls the study	Scientific method	Rationalistic/ empiricist origins Causality can be established Quantitative methods
Postpositivist	A knowable reality that is theory-bound	Observable, but not necessarily through rigorous application of the scientific method Observation dependent upon theoretical context Researcher directs and interprets	Primarily quasi-experimental	Primarily quantitative methods but qualitative methods can be used
Constructivist	Multiple socially constructed realities which may change throughout the study	The researcher and the subjects interact and influence each other Hermeneutics: interpretive understanding/meaning Researcher directs and interprets but values must be clearly defined as they impact upon interpretation	Interviews Observations Document reviews	Rejection of objective reality Reality is confirmed through • tracking data to its source • extended and substantial engagement and observation and through triangulation of • multiple data sources • multiple methods of data collection

Table 2.1 cont.: Research paradigms

Paradigm	Ontology	Epistemology	Methodology	Notes
Transformative	Multiple realities arising particularly from social, political, cultural, economic, ethnic, gender and disability values Reality construction to be examined in the light of oppressive influences	Interaction between the researcher and the researched Empowerment of the researched	Interviews Observations Document reviews Shift in power from the researcher to the researched Inclusion of diverse ‘voices’ in the analysis	Includes critical theorists, participatory action researchers, and disadvantaged minority groups Directly addresses political questions Quantitative, qualitative and mixed-method approaches Avoidance of bias of particular import
Pragmatic	Ignores concept of reality and focuses on ‘what works’ within the defined frame of reference of the study	Interaction between the researcher and the researched Contextualised/socially or historically situated researcher	Method selected for purpose	Provides underlying philosophical framework for mixed-methods approaches Emphasis on common sense and practical thinking Effectiveness is the yardstick for value – results count

Summarised from Mertens, 2005, pp. 8-29

2.1 Marriage of convenience – evidence from practice

Supportive evidence for the implementation of an educational strategy can only be found in the classroom. Some might argue that ‘improved grades’ offer indicators of improved learning, but there is more here than meets the eye! Evidence of student learning needs to be examined in the light of the following:

- a clearly articulated teacher’s definition of student learning;
- an understanding of the student’s perception of their learning;
- common threads between the above;
- methods of data collection which enable detection of the desired learning;
- clear documentation of the action involved in teaching, learning and assessment;
- strategies that align teacher expectations, student performances and assessment of those performances.

2.1.1 Why not experiment?

The researcher’s experience has identified the university classroom as a dynamic and interactive system in which the teacher has limited control over impacting variables. The complex fabric of culture, history and experience that has led the mix of students to the classroom is accompanied by an almost infinite range of resultant impacting variables that inherently defy experimental control (Alexander and Hedberg, 1994). Even if it were attempted to account for such differences through some randomised design, its external validity would be questionable to say the least (Alexander and Hedberg, 1994). The following questions need then to be asked:

1. Would useful resolutions to complex social problems/questions result from studies that preclude any of the social dynamics that intrinsically drive the social interactions?
2. Can real solutions even exist in such a social vacuum?
3. Is it ethical to deprive some students and imbue others with a teaching/learning framework that is believed would be advantageous to all?

Affirmative responses to some or all of these questions might prove untenable to a committed teacher. Indeed the following viewpoints underpin the teacher's attitudes to such an approach:

1. Pervading this study, has been the teacher/researcher's belief that complex social interactions drive student learning, expanding their cultural, moral and historical perspectives, and enhancing their opportunities to learn by confronting them with multiple perspectives of the questions, problems or tasks they undertake;
2. In a traditional, expository learning environment, a social vacuum might present no problem. However such solutions offer no real advantage to students that would fit them for global communities and their demands for lifelong learning;
3. To deprive students of what is perceived as an advantageous educational intervention seems to challenge a prime aim of a teacher's professional practice – to develop each student to their potential.

Indeed within the context of qualitative studies, it remains incumbent upon researchers to maintain an ethical regard for the human faces of their studies.

> Should the quest for objectivity supersede the human side of those whom we study? … Clearly as we move forward with sociology, we cannot … let the methods dictate our images of human beings. As Punch (1986) suggests, as field-workers we need to exercise common sense and moral responsibility, and we would like to add, to our subjects first, to the study next, and to ourselves last. (Fontana and Frey, 1994, p. 373)

A focus of this study has been the improvement of student learning, and student actions and reactions have consequently figured prominently throughout the implementations and in its evaluation. This qualitative study has recognised the bounded, complex social system of the classroom and relied upon a stringent evaluation protocol to examine its action.

2.1.2 Mixed methods

> Pragmatists decide on what they want to research, guided by their personal value systems; that is, they study what they think is important to study. They

> then study the topic in a way that is congruent with their value system, including variables, and units of analysis that they feel are most appropriate for finding an answer to their research question. (Tashakkori and Teddlie, 1998, p.27, cited in Mertens, 2005, p. 295)

This study has been driven by action in the classroom. The research paradigm and methodology have been forced by the selection of the specific teaching/learning contexts. So too has been the decision to use a mixed methods approach to collecting data. Evidentiary sources include both quantitative data (assessment) and qualitative data (surveys, observations etc). Analysis of these data has afforded evidence of multiple perspectives of student learning. Hence, parallel, but independent, collection of data was selected (Mertens, 2005, Cresswell, 2002). Triangulation of all of the data sources (identifying repeating themes) gives greater breadth of understanding (Cresswell, 2002). Analysis can afford opportunity for 'meta-inferences' based on the two types of data (Mertens, 2005), and comparisons can be used to identify common ground (Cresswell, 2002). In this study: comparison of ranks in achievement (assessment data) and perceived learning (student surveys) was used to test the contention that students *know what they know* and *know what they do not know.*.

2.2 Case studies

> Exploration of a bounded system is often referred to as a case study. … (This) method of studying elements of the social through comprehensive description and analysis of a single situation or case, for example, a detailed study of an individual, group, episode, event, or any other unit of social life organisation. Emphasis is often placed on understanding the unity and wholeness of the particular case. (O'Leary, 2005, p. 115)

O'Leary (2005) has questioned the inclusion of 'case studies' as a methodology as they may use multiple data sources and access a number of other methodologies. Chapter 6 of this thesis has been presented as a case study exploring teaching, learning and assessment across multiple sessions in a fundamental statistics subject. The methodology of action research has been applied to the study of the development of appropriate learning strategies addressing *statistical thinking*, and the evaluation process has included comparisons across the multiple implementations. Chapter 7 presents a further case study, a single implementation of a teaching/learning framework targeted at development of *critical and evaluative thinking* in accounting students.

Case studies have also been described by Mertens (2005) as qualitative research strategies rather than methods, as a variety of approaches may be used for data collection. Commonality in definitions recognises case studies as investigations which focus on a single instance in a single context. Description, analysis and interpretation provide particular understanding of the instance in context. Generalisations beyond the confines of each study have not necessarily been the primary foci of the studies. Their main purpose has resided in the development and evaluation of two differing learning environments. However, case studies may also be of interest for their potential to provide more global insight. This can be achieved through collective comparison and contrast of the results of similar studies in order to identify common conclusions. (O'Leary, 2005; Stake, 1995).

2.3 Selected methodologies

The researcher's own practice of her profession has been hauled in for scrutiny in this study. Her ever-present attention to classroom change in the learning of her students has been heightened throughout. Early deliberations led to the identification of all potential sources of evidence of student learning, and well-springs of reflective comment on successes and shortfalls and the reasons behind them. The combination of the researcher's adherence to social constructivism in her classroom, and the potential sources of data for this study have precipitated the qualitative methodological choices (see Table 2.2).

The research designs are displayed in Table 2.3. The specified selections of methodologies are discussed in the ensuing paragraphs to further elucidate connection between practice and theory.

Table 2.2: Choices of Methodologies for the classroom studies

Case studies*	Methodology	Data collection and analysis
Statistical learning (Chapter 6)	Grounded Action Phenomenology Ethnography	Surveys Assessment data Deconstruction of assessment tasks Peer discussion Comparisons across multiple implementations Researcher's journal annotated by involved parties Reflective Practice
Development of critical thinking (Chapter 7)	Grounded Phenomenology Ethnography	Interviews Surveys Assessment data Peer discussion Researcher's journal annotated by involved parties Reflective Practice
Assessing knowledge and skills (Chapters 6 and 7)	Grounded Action Ethnography	Surveys Assessment data Deconstruction of the cognitive demand of assessment through application of a learning taxonomy Peer discussion Researcher's journal annotated by involved parties Reflective Practice

* Not strictly a methodology, but used in conjunction with the accompanying methodologies. See paragraph 2.2.

Table 2.3: Research design

See print copy for Table 2.3

* Not strictly a methodology, but used in conjunction with the accompanying methodologies See paragraph 2.2.

Adapted from O'Leary, 2005, p. 100

2.3.1 Action research

> *Improvement* and *involvement* are central to action research. There is, first, the improvement of a *practice* of some kind; second, the improvement of the *understanding* of a practice by its practitioners; and third, the improvement of the *situation* in which the practice takes place. (Robson, 2002, p.215)

As with experiential learning theory, action research has been heavily influenced by Lewin and his work in institutional change within large organisations. Action research addresses practical problems and seeks solutions to those problems in context (O'Leary, 2005; Cresswell, 2002; Robson, 2002). It affords opportunity for promoting change in professional practice. For this reason, it provides a relevant strategy for examining social programs in general, and educational programs in particular. Gomm (2004), however, has regarded such research by participant researchers in professional practice as 'elitist' attempts to raise the profile of the associated profession and has further claimed that it frequently generates unconvincing justifications for resultant conclusions. Green (2003) has also cautioned against the participant role of a researcher, highlighting the potential for value laden conclusions. However she has also anticipated that the solution to this dilemma might lie in a well defined evaluation program which addresses these issues.

In action research, knowledge and change are inextricably linked in that the cycle incorporates knowledge that fosters change that generates more knowledge and so on. (O'Leary, 2005). This is part and parcel of the daily practice of the teacher whose actions are continually reformed by student reactions and reflection upon the process (Cresswell, 2005). In a well designed constructive learning environment, the teacher and the learners are collaborators in the learning journey. So extension of this existent cyclic process to include research, meets another defining characteristic of action research (Cresswell, 2002).

A prime challenge for the researcher, then, is to devise a systematic process of evaluation that includes strict accountability for influencing values, the action and all evidentiary sources. This is especially true in the case of the participant researcher. Ensuring effective negotiation with all stakeholders may also present challenges, but must be achieved if the goals of the research are not to be diverted by conflicting values (O'Leary, 2005).

Action research may take the form of a spiralling chain of cycles (Figure 2.3). The first phase involves a period of reflective observation and design. The designed interventions feed into the first implementation and each cyclic implementation is followed by analysis, reflection and further design. The results of each cycle of deliberations feed into each subsequent implementation.

O'Leary (2005) includes action in her first cycle. But in the strategy illustrated in Figure 2.3, the first phase only includes observation, critical evaluation and planning. The results of this phase are implemented in the first action phase. As a consequence, O'Leary begins each phase with observation whereas Figure 2.3 reflects the researcher's perception of deliberate action followed by reflection exemplified in her teaching.

2.3.2 Phenomenology

> Phenomenology is premised on a world that is: (1) 'constructed' – people are creative agents in building a social world; and (2) 'intersubjective' – we experience the world with and through others. … (Phenomenology is the study) of phenomena as they present themselves in direct experience. (O'Leary, 2005, p. 122)

The foci of phenomenological studies are the 'rich' reported descriptions (or artistic depictions) of the contextual experiences of individuals. Hence their personal constructs are of prime importance, independent of any perception of 'reality'. Meanings arise throughout the collection, analysis and recording of their descriptions. The 'story teller' and the subject collaborate in the 'story writing'. (O'Leary, 2005)

Examining the impacts upon the protagonists in a complex classroom environment may require exploration of individual experiences of the structures and processes involved, and 'rich' descriptions of their involvement may provide evidence for evaluating the action. In this research, deconstruction of the teaching, tutoring/marking and student experiences have revealed multi-faceted images of the implemented learning frameworks.

Please see print copy for Figure 2.3

Figure 2.3: Action research spiral

Adapted from O’Leary, 2005, p. 141

2.3.3 Ethnography

> Ethnography can be described as a research method designed to describe and analyse practices and beliefs of cultures and communities (Patton, 2002). A key assumption is that by entering into firsthand interaction with people in their everyday lives, ethnographers can reach a better understanding of the beliefs, motivations, and behaviours of the people in the study than could be achieved by any other method. (Mertens, 2005, p. 235)

The aim of ethnographic studies is to develop understanding from the 'inside'. This requires background study of the cultures under examination. Because of the immersion of the reporting researcher in this study, there is some scope for criticism of her lack of objectivity. Whilst many ethnographic studies explore the crossing of *cultural boundaries*, the elements of ethnography that impact upon this study, involve the cultures of 'student' and 'teacher' and, at times, the confrontation between those two *cultures*.

The Advanced Oxford Dictionary online (2005) defines culture as

> … the customs and beliefs, art, way of life and social organization of a particular country or group. (Oxford Advanced Learner's Dictionary, 2005)

In this study, the researcher encountered conflicts in and between the cultures of 'teacher' and 'student'. The 'teacher' culture was embedded in collegial, discipline, research and faculty histories and demands. Similarly the 'student' culture was heavily influenced by age (maturational effects), academic histories and perceptions, and work, family, and university demands. Their reactions to the complex teaching environment drove many of the changes as the project progressed. They also fuelled much introspection by the participating researcher/teacher, introducing elements of *auto ethnography* to the study (Denzin and Lincoln, 2003).

2.3.4 Grounded Research

Grounded theory is a methodology which can generate theory from data. The initial phase involves the definition of the question/topic and the design of the data collection protocols. Subsequent phases require the analysis and interpretation of the data based upon the theory, and may involve collection of further data. (O'Leary, 2005; Cresswell, 2002) Thus

> … grounded theory researchers know from the planning phase of their study that much of *their* methodological protocol cannot be developed in advance and is indeed dependent on what emerges from the initial data. …Grounded theory may be flexible, iterative, and emergent, but it is never ill-defined, haphazard, or *ad hoc.* (O'Leary, 2005, p. 98)

2.3.5 Evaluation research

It is important to devise evaluation standards in order to operationalise the variables under examination. (Gomm, 2004) The case study reported in Chapter 6, tracks the learning of statistics students. Detection of the 'learning' required:

- clarification of the researcher's perception of statistical learning;
- definition of required knowledge and skills by discipline 'experts';
- quantification of student learning through achievement records (marks, grades, student perceptions of learning) and student behaviours (attendance and submission patterns); and
- deconstruction of the assessment and marking criteria in terms of the required knowledge and skills.

In undertaking such detection, the researcher recognised a plurality of perceptions of learning that may even conflict. However, the attitudes and efforts of all interested and involved parties merited consideration in the devising of successful strategies that offered optimum classroom and discipline learning, and also fostered the development of deeper, independent and lifelong approaches to learning.

Table 2.4: Addressing criteria for judging the quality of the qualitative research in this study

Criteria	Evidence
Credibility (parallels internal validity)	
• Prolonged, substantial engagement	• Multiple implementations in the statistical learning case study.
• Persistent observation	• Researcher's journal and annotations of interested parties and peer reviewed publications.
• Peer debriefing	• Researcher's journal and annotations of interested parties and peer reviewed publications.
• Negative case analysis	• Some comparisons were possible between the observation phase and implementation phases in Chapter 6; some data comparisons were made between before and after implementation in Chapter 7.
• Documentation of progressive subjectivity	• Researcher's journal and annotations of interested parties.
• Continuous corroboration with respondents/participants	• Surveys, assessment data, researcher's journal and annotations of interested parties.
• Triangulation (consistency of evidence across multiple sources of data)	• Surveys, assessment data, commonalities in two case studies.
Transferability (parallels external validity)	
• Extensive description of context and culture	• Researcher's journal and annotations of interested parties and peer reviewed publications.
• Multiple cases (incorporating the relationship to relevant theory)	• Commonalities in two case studies.
Dependability (parallels reliability)	
• Changes should tracked and the research process publicly documented	• Researcher's journal and annotations of interested parties and peer reviewed publications.
Confirmability (parallels objectivity)	

Table 2.4: Addressing criteria for judging the quality of the qualitative research in this study

Criteria	Evidence
• Data can be tracked to its source and deductions should be explicit	• Case studies in Chapters 6 and 7.
Authenticity	
• Fair (Conflicts and value differences should be overt)	• Researcher's reflective comments and discussions.
• Ontological: Documentation of changes in individual and group constructions	• Researcher's journal and annotations of interested parties, peer reviewed publications and assessment data.
• Catalytic: the extent to which the study drives change	• Case studies in Chapters 6 & 7.
Emancipatory	
• Epistemological: Declared position of the researcher	• Chapters 2, 3 and 4
• Linkage to community improvement	• Peer reviewed publications, student surveys
• Giving voice to the silent and the marginalised	• Few students were identified as 'zero fail' in the case studies and no further examination of these students was conducted.
• Critically reflective and personally aware	• Researcher's reflective comments and discussions
• Engages the trust of the participants	• Participant information sheets and interactive engagement in the classroom
• Sharing of wealth generated by the research.	• This was not identified as relevant in this research

Adapted from Mertens, 2005, pp. 253-259 and from Guba and Lincoln, 1989, pp. 233-250.

2.4 Participant researcher

Adler and Adler (1994) categorise the participation of qualitative researchers in the following way:

Peripheral-member-researcher	Observing and interacting sufficiently to establish an insider's perspective without participating in the activities of the core group
Active-member-researcher	More involved in the groups activities, but not accepting their values or goals
Complete-member-researcher	Becoming an active member throughout the study

(Adapted from Adler and Adler, 1994, p. 379-380)

As a teacher involved in the teaching of statistics in the primary case study, the researcher was appropriately positioned to undertake that study as a 'complete-member-researcher' (Adler and Adler, 1994). However in the second case study, the researcher was a facilitator for the participant teachers and hence might at best be described in terms of classroom teaching and observation as a peripheral-member-researcher. However, for this case study, the researcher facilitated the subject design, and hence has declared an active involvement in the discussions and evaluations involving participating staff.

Mertens (2005) recommends that the participant researcher is positioned to observe:

- the physical environment;
- the human and social environment;
- the program activities and participant behaviours;
- informal interactions and unplanned activities
- immersion in the language and 'culture' of the participants (in this study, the researcher refers here to the 'culture' of students encompassing the demands upon them and their coping strategies in particular);
- nonverbal communication;
- unobtrusive measures;

- what does not happen (non attendance (attendance rates), inactivity, lack of completion/submission, lack of questioning, plagiarism…).

As a teacher of some experience, focused on the learning of her students, such observations form a part of her regular teaching practice (although rarely formally documented) in interacting with students in the classroom.

Gomm (2004) expressed reservations about practitioners as researchers. He advised that project evaluation should allow the researcher to ascribe observed results to the intervention/s under examination. He claimed that the lack of research design implementing control over other potentially influencing variables renders many arguments for cause/effect unconvincing. However other authors such as Mertens (2005) have argued that although causality belongs strictly in the realm of the experimentalist, judgement of the merits of positivist, quantitative research has its parallel for qualitative research (Guba and Lincoln, 1989). Guba and Lincoln's criteria for such judgement have been discussed in Mertens (2005) and have been related to this study in Table 2.4.

2.5 Ethics

This study has involved the tracking of student learning in two undergraduate subjects at universities in Australia. Privacy of students has been accorded the highest consideration. All survey and permission forms and survey files containing identifying information have been maintained under lock and key, accessible only to the researcher and her supervisors. All participating students were over eighteen years of age and informed consent was sought from them prior to collection of the data. The study involved no invasive treatments (beyond exposure to statistical learning (Chapter 6) and practice in critical evaluation of accounting theories and philosophies (Chapter 7)), and the data collected from the students required:

- self-completed surveys;
- assessment data (with all identifying detail removed) from college assessment files;
- comments reported in the researcher's journal and annotated by collaborating parties.

Teachers, tutors and markers were similarly surveyed and peer discussions also reported in the researcher's journal.

The primary case study was conducted primarily at the University of Wollongong during the researcher's candidature for her Doctorate (see Chapter 6). The secondary study represents an implementation of related strategies at the University of Western Sydney (see Chapter 7) during the same period. The researcher was a participant teacher in one of the case studies and co-designed the learning framework for the second. The researcher has declared her interest in the studies and has overtly discussed any 'values'/biases resulting from her espoused theories of teaching, learning and assessment.

Clearance for the conduct of this study was sought and granted through the University of Wollongong (www.uow.edu.au/research/rso/ethics/human).

2.6 Limitations

The following considerations may highlight limitations in comparability of data across implementations in the disparate case studies of this thesis:

- Each of the case studies undertaken in this project was considered as a closed system. The aim in studying each learning environment was perceived as of intrinsic value to teaching and learning in its corresponding discipline. Multiple implementations and/or commonalities in case studies have enabled some generalised comment;
- There was some variation in the cohorts enrolled in STAT131 (see Chapter 6) between sessions (implementations);
- There was considerable variation in experience and expertise in the tutoring staff between sessions (implementations) and in the secondary case study. This was particularly evident in their commitment to enlisting student feedback through the surveys. Hence response rates have varied;
- Marking underwent substantial development throughout the studies;
- Student surveys were completed online in STAT131 (see Chapter 6) and although most students completed the surveys during class time, some did not. Although student access to the surveys was password protected, this

does not preclude the possibility that the surveys may have been completed by non participants in the subject who were known to participants;

- Survey completion for the second case study (see Chapter 7) was by hand, in class, at the end of the session. Some students may have submitted trivial responses so that they could leave early from class;
- Although most survey questions offered a selection of responses, student honesty cannot be guaranteed;
- Analyses have identified a significant increase in cognitive demand of the assessment across implementations.

Chapter 3
Learning: Now you see it … or do you?

… "Are there any lions or tigers about here?" she asked timidly.

"It's only the Red King snoring," said Tweedledee. …

"He's dreaming now, " said Tweedledee:"and what do you think he's dreaming about?"

Alice said "Nobody can guess that."

"Why, about you!" Tweedledee exclaimed, clapping his hands triumphantly. "And if he left off dreaming about you, where do you suppose you'd be?"

"Where I am now, of course," said Alice.

"Not you!" Tweedledee retorted contemptuously. "You'd be nowhere. Why, you're only a sort of thing in his dream!"

"If that there King was to wake," added Tweedledum, "you'd go out – bang! – just like a candle!"

"I shouldn't!" Alice exclaimed indignantly. "Besides, if *I'm* only a sort of thing in his dream, what are *you* I should like to know?"

"Ditto," said Tweedledum.

"Ditto, ditto," cried Tweedledee.

He shouted this so loud that Alice couldn't help saying, "Hush! You'll be waking him, I'm afraid, if you make so much noise."

"Well it's no use *your* talking about waking him," said Tweedledum, "when you're only one of the things in his dream. You know very well you're not real."

"I *am* real!" said Alice and began to cry.

"You won't make yourself a bit realer by crying," Tweedledee remarked: "there's nothing to cry about."

"If I wasn't real," Alice said – half-laughing through her tears, it all seemed so ridiculous – "I shouldn't be able to cry."

"I hope you don't suppose those are *real* tears?" Tweedledum interrupted in a tone of great contempt.

From "Tweedledum and Tweedledee" in
Through the Looking Glass by Lewis Carroll, pp. 161-163.

3.0 What is learning?

> "…the way in which an educator builds his curriculum, selects his materials, and chooses his instructional techniques depends, to a large degree upon how he (she) defines "learning". Hence, a theory of learning may function as an analytical tool; its exponents can use it to judge the quality of a particular classroom situation." (Bigge, 1976, p.4)

In the dynamic, interactive and multi-dimensional communication structure of the classroom, it is relevant to consider what student learning is, and to examine how it is accomplished. With so many potentially influential factors, it is also essential to determine how the learning can be best guided and ultimately how the teacher might recognise that it has in fact taken place.

Watkins et al. (2002) identify the three most common modes of discussing learning:

1. S/he taught me … [Instruction]
2. I made sense of … [Construction]
3. We worked out that … [Co-construction] (Watkins et al., 2002, p.3)

The *ways of learning* are contingent upon the teaching/learning frameworks that foster them (a focus in Chapter 4). Student engagement in the classroom activities will reflect the influences of the learning theories underpinning those frameworks. Similar practices may be dealt with differently by different theories (Watkins et al., 2002) and accents on roles of the protagonists will vary.

3.1 Learning theories

In this chapter, the fundamental precepts of several relevant learning theories are pursued, potential classifications of learning surveyed, and other aspects of learning that might link with this study are identified.

As the researcher tracked an historical perspective of several of the main learning theories, she identified salient aspects of several different theories. No one approach satisfied her understanding of student learning and its promotion. It became apparent that some theorists whose research spanned the rise (and fall) of the popular persuasions, shifted their ideological leanings as their knowledge-base expanded.

Parallel to the developmental history of learning theories have been changes in the methodological approaches to research in the social sciences. In positivist traditions, the researcher is viewed as independent of the subject of study; evidence is derived solely from the observable (Mertens, 2005). Behaviourism arose in this tradition. Behaviourists held the researcher as independent and in control of the learning environment and believed that learning results from the formation of associations. Evidence of learning was provided by changes in the observable behaviour of the learner (Gredler, 2001).

Postpositivist researchers still sought objective observation of learning, but acknowledged that observations were influenced by the researcher's background knowledge and beliefs (Mertens, 2005). Postpositivist approaches to learning theory acknowledged an internalisation of the learning as the learner constructed meaning, but that this process could not be readily observed. There was, however, a shift in emphasis away from a controlled learning environment to one centred on the learner and his/her interactions with that environment. (Gredler, 2001)

Constructivists not only acknowledged the impact of the researcher's values and beliefs, but recognised interaction between the researcher and the researched and the social construction of knowledge. They sought interpretive understanding in context rather than an independent observed reality (Mertens, 2005).

3.1.1 Behaviourism

Perhaps the earliest of the modern learning theories is that rooted in the scientific experimental beginnings of psychology. Behaviourist investigations focused on detached observations of the behaviours they believed to be manifestations of internal (and unobservable) mental functioning resulting from an instructional process. They viewed learning as the process of forming associations between basic elements, but their attention was centred on the behaviours. Much of the theory arose from experimental investigation of animal behaviour in which the researcher controlled the environment.(Gredler, 2001: McInerney and McInerney, 1994)

Watson (1878-1958) is credited with the term *behaviourism*. He was inspired by the work of Pavlov (1849-1936) and his conditioning experiments in learning centred on the stimulus/response mechanism of a hungry dog. Watson applied the underlying theory to education and developed it further, introducing the notions of reinforcement (both positive and negative rewards) for the strengthening of responses.

Skinner (1904-1990) drew also upon animal training techniques to develop his principles of operant conditioning. Classroom application viewed students in a strictly controlled environment with teaching revolving around sequenced steps toward achievement of a clearly defined learning goal. Successful completion of each step was reinforced and the student then moved to each subsequent step until the goal of the instruction was finally achieved. Success at each stage was essential for reinforcement and promotion to the next. Motivation to learn was independent of the learner and entirely in the hands of the reinforcer, the teacher.

The most pervasive legacy of the behaviourists lies in assessment. There, evaluation of student learning relies heavily on observable behaviours indicative of achievement of teaching objectives and evidenced in examinations and other tasks. However, aspects of affective knowledge such as ‘preparedness’ or ‘acquisition of values’ are not readily demonstrable and might even prove ethically beyond the bounds of classroom assessment. Other behaviours, reflected in student reactions to the teaching environment, might also be productively used by the teacher to orchestrate the teaching process of initiation, rephrase and revision.

Critics of *behaviourism* have claimed the control of the learning to be in the hands of the teacher, with the student behaviour manipulated through the use of reinforcement. This has been perceived as restrictive of the student’s personal construction of knowledge. (McInerney and McInerney, 1994) The ‘rewards’ potentially provide only the extrinsic motivation that is frequently associated with *shallow learning.* Skinner’s concepts of reinforcement have not been perceived as adequately explaining intrinsic motivation (Bruner, 2000; Driscoll, 2000) which is strongly aligned with greater motivation to learn

> … students show greater motivation when they have an internal, as opposed to an external, orientation. This means that they tend to perceive learning tasks as skill determined and thus subject to personal control. Externally oriented students tend to believe that their success at a learning task will be determined by chance rather than by means within their control. These students are therefore less likely to be motivated to engage in the learning task. (Driscoll, 2000, p. 304-305)

There has been much recent debate over the potential of extrinsic motivation to diminish intrinsic motivation. Cameron et al. (2005) and Pintrich et al., (1990) have claimed that extrinsic motivation may in appropriate circumstances actually enhance intrinsic motivation. Bruner (2000), on the other hand, argued that extrinsic motivators may drive less sophisticated learning, resulting in the development of

> … rote abilities and depend on being able to "give back" what is expected … (Bruner, 2000, p. 221)

He claimed that 'discovery learning' has the effect of promoting intrinsic motivation

> … to the degree that one is able to approach learning as a task of discovering something rather than "learning about" it, to that degree will there be a tendency for the child to carry out his learning activities with the autonomy of self-reward or, more properly by reward that is discovery itself. (Bruner, 2000, p.221)

He contended that

> When learning in the short run leads only to pellets of this or that rather than to mastery in the long run, then behaviour can be readily "shaped" by extrinsic rewards. When behaviour becomes more long-range and competence-oriented, it comes under the control of more complex cognitive structures, plans and the like, and operates more from the inside out. (Bruner, 2000, p.223)

Such debate, however, should caution unconsidered use of extrinsic motivation in the university setting that seeks to promote *lifelong learning*.

Gestalt psychologists reacted against the constrictions of behaviourism. Advocates were concerned with the more cognitive aspects of learning and addressed problem solving and associated thinking (Gredler, 2001). They identified the difference between sensory experience and an individual's perception of that experience. The latter arose from cognitive modification of stimuli based upon prior experience or knowledge. The Gestalt movement thus viewed knowledge as arising not merely from experience, but also from active reorganisation of that experience. They recognised 'insight' as an ability to assemble disparate elements into a solution to a problem, for example: the act of constructing a tool. Whilst the behaviourist teacher might focus on explicit behaviours and assessment which acknowledges exhibition of those behaviours, for the Gestalt school, learning only resulted when learner performance was accompanied by some realisation of its importance. (McInerney and McInerney, 1994)

A burgeoning recognition of the necessity to return the study of learning to the classroom led to the emergence of models of instructional theory. In the behaviourist camp, Skinner described his model of programmed instruction. A more cognitive focus was provided by Piaget's model which regarded cognitive development as based upon physical and social maturation.

An alternative viewpoint was developed by Bruner. His model of cognitive development was founded on a hierarchy of states of 'readiness'. Bruner attempted to develop a theory of instruction based upon the developmental theory of Piaget. However, he observed wide variation in ages of attainment of Piaget's various stages - too wide to be of use in the classroom. The notion of sequencing underpinned his instructional theory. Ultimately he described stages in terms of observable indicators of levels of development rather than age bands: enactive (use of concrete learning strategies for new concepts); iconic (use of images/diagrams to represent concepts/procedures) and symbolic (use of familiar symbols and language to teach new concepts). He proposed construction of a spiral curriculum to enable entry to instruction at the most appropriate level. (Bruner, 1971)

3.1.2 Cognitive learning theory

World War II provided a major impetus for change in the development of learning theory. It introduced the need for new training tactics to address complex learning requiring

> … the perception, judgement, and decision-making process of the human operators. (Gredler, 2001)

The advent of high speed computers also inspired learning theorists, giving rise to fresh models of human thought. Schools of learning arose from these perspectives. Information processing models compared learning to the internal functioning of the computer. Terms such as:

- 'input', 'process' and 'output';
- 'short' and 'long term memory'; and
- 'encoding', 'retrieval' and 'chunking'

were coined to describe the passage from instruction to learning. (Gredler, 2001)

Cognitivists regarded learning as an internalised process in which the learner constructed new knowledge. (McInerney and McInerney, 1994) This process was both impersonal and objective, and the learner dispassionately acquired new knowledge and skills logically. Such a definition of learning presented a fresh dilemma for research: experimentally determined evidence of learning was more problematic as it was not apparent and could only be detected in observable manifestations of that learning or through the learner's reporting of the processes. Educational research expanded into the qualitative arena of qualitative research and came to include action research, phenomenology, ethnography, grounded research and others (Mertens, 2005).

Meaningful learning founded on 'schema theory' (which afforded organisation of prior knowledge) was described by Ausubel (1918-). The notion of organisational structures underpinned his theories. He advocated introducing structures (*organisers)* which systematised knowledge and facilitated assimilation of new knowledge and skills by providing *ideational scaffolding*. (Ausubel, 1960, Ausubel, 1968, Ausubel, 1978) These notions have been found useful to the researcher in the classroom and have been discussed more fully in Chapter 4.

Important to this study are some of the theories incorporating aspects of interactive learning. Those which influenced the teaching of the researcher have been singled out for inclusion here and these too have been discussed more fully in Chapter 4.

3.1.3 Constructivist theory

Constructivists varied in the emphases placed upon social, cultural and personal aspect of learning. Teaching approaches were affected by adherence to particular theoretical emphases: external factors (modelling, and explanation); interpersonal (collaboration between all protagonists in the classroom); and intrapersonal (learner-driven, discovery/exploration). Many cognitive theorists (such as Vygotsky, Bruner and Ausubel) developed towards a more 'social constructivist' perspective throughout their careers.

Although Bruner's theories altered throughout his life, a fundamental concept permeating his research was the notion of learning as a meaningful *organisation* of an individual's view of the world. He became increasingly interested in social and cultural impacts upon education. Inspired by the work of Vygotsky, he began to examine the interactive social environment in which students come to construct meaning. He embraced 'discovery learning' which focused on student interactive learning experiences.

In contrast to the developmental theory of Piaget, Vygotsky (1896-1934) believed that human biological development could not be isolated from social and cultural influences upon cognitive growth. Themes permeating his work include:

- higher mental functioning which included the internalisation of mental interaction with others;
- language acquisition, from learning words and thinking aloud to inner speech associated with higher mental functioning;
- mediation of learning through the agency of the teacher;
- culture which provides the rich fabric of individual experience as background to the learning.

He observed student learning in what he referred to as the 'zone of proximal development', a term he coined to describe the period when students have not yet acquired the knowledge/skill and require expert assistance. This phase he described as bridging the gap between current and potential development and whilst the lower limit might be bounded by prior achievement, he believed the upper limit to be a function of the social learning environment encompassing the teacher, structured instruction, peers and the physical environment. Thus he claimed that a student's potential development is augmented by the pedagogical design and opportunities for collaboration during the learning process. (Daniels, 2003, Ussher and Gibbes, 2002)

3.1.4 And deeper still: Co-constructivist theory

Co-constructivist theory expands the constructivist philosophy's address of the cognitive and social impacts upon learning to emphasise the view of students and teachers as collaborators. Vygotsky's view of the teacher as collaborating through mediation permits acknowledgement of him as a co-constructivist. (Ussher and Gibbes, 2002)

The focus in a constructivist classroom is the individual who seeks to interpret his/her environment through the process of learning. Co-constructivism removes the focus from individual cognitive construction to include the teacher and peers as collaborators. Responsibility for learning is thus shared.

Learner disposition can produce positive behaviours (those affirming intrinsic motivation, success as a result of effort and problem solving approaches to challenge) or negative behaviours (those affirming competitive values, success as a result of ability and poor self worth in the face of challenge). A rich learning environment founded on experience and collaboration is more likely to induce positive learner attitudes than more traditional approaches. (Carnell, 2007; Klenowski et al., 2006; Watkins et al., 2002)

Please see print copy for Figure 3.1

Figure 3.1: Expanding Kolb's learning cycle to include learning about learning
(Watkins et al., 2002, p. 6)

Effective learning requires exposure to

- active participation (*doing*);
- reflective observation (*reviewing*);
- accommodation of theory and experience (*learning*) and
- testing the learning (*applying*).

This represents a simplified version of Kolb's learning cycle (see Figure 3.2). Watkins et al. (2002) overlayed a further cycle addressing *learning to learn* (see figure

3.1). Promotion of such *meta-learning* is believed to increase independence and foster *lifelong learning.* (Watkins et al., 2002)

3.1.5 Experiential learning

> … this learning process must be reimbued with the texture and feeling of human experiences shared and interpreted through dialogue with one another. … We lost touch with our own experience as the source personal learning and development … (Kolb, 1984, p.2)

Kolb's work found its roots in the educational theories developed by Dewey, Lewin and Piaget. Dewey developed experiential learning programs for higher education. These programs increasingly conflicted with the more expositional approaches of traditional education. He recognised that the challenge was to meet the needs of *lifelong learning* demanded by growing changes in the workplace and society in general. He advocated reinstatement of

> … apprenticeships, internships, work/study programs, cooperative education, studio arts, laboratory studies and field projects. (Kolb, 1984, p. 5)

The methods all demanded

> … the learner is directly in touch with the realities being studied … (Keeton and Tate, 1978, p. 2 in Kolb, 1984, p. 5)

Emphasis on *lifelong learning* and the entry of students whose experiential background rather than academic achievement fulfilled university entrance criteria (West, 1998) has provided added impetus to review of the merits of experiential learning techniques. Similarly there has been a growing call for justification for traditional curricula rather than the knowledge bases acquired by mature learners with a breadth of experience acquired in practice. Such approaches afford opportunities to achieve for those students who, although perhaps wanting in academic background, have demonstrated their successful learning through active and successful engagement in the workforce. (Kolb, 1984)

The marriage of learning through experience and accreditation founded on reliable assessment against competency-based standards or outcomes, promotes the perception of higher education as more than exposition of its vast reservoirs of knowledge. A more holistic approach to curriculum building presents the challenge to not only *fit* students for the workplace, but for also community and life in general (West, 1998).

Lewin's description of action research methodology has proved constructive in the study of functioning groups ranging from small social groups to large and complete organisations. Lewin believed that the conflicts arising between conceptual theory and immediate and subjective experience facilitated learning. In contrast to the behaviourist school and classic definitions of knowledge gained through detached observation, his focus was on subjective learning. He claimed that this focus

> … has developed into a strong commitment in the practice of experiential learning to existential values of personal involvement, and responsibility and humanistic values … (Kolb, 1984, p. 10)

In Piaget's developmental theory, Kolb perceived a description of

> … how intelligence is shaped by experience … (it) arises as a product of the interaction between the person and his or her environment. (Kolb, 1984, p. 12)

Similarly inspired by Piaget, Bruner forged the link between cognitive theory and teaching practice in his theory of instruction. His instructional theory afforded curricula appropriately targeted at all stages of cognitive development. Experience-based programs translated the conceptually complex detail of mathematics and science into forms more aligned with learners at more concrete stages of development. They also promoted more individualised and self-directed approaches.

Kolb's model viewed learning as a continuous process across two dimensions:

- the processing dimension: a continuum spanning from *active experimentation* to *reflective observation* and
- the perception dimension: a continuum spanning from *concrete experience to abstract conceptualisation.*

He believed the most effective learning experiences result from opportunities to experience all four phases of his learning cycle (illustrated in Figure 3.2):

1. Phase 1: concrete experience and active experimentation. This phase challenged involvement in new experiences in the light of prior knowledge;
2. Phase 2: concrete experimentation and reflective observation. Learners interpret experience providing explanation of the preferred options;
3. Phase 3: abstract conceptualisation and reflective observation. Connections are made between the preferred options and explanations constructed with reference to theory. Thus theory provides external justification;
4. Phase 4: abstract conceptualisation and active experimentation. The learner tests concepts and skills in real situations.

Although recommending that learners should be exposed to opportunities for all four phases, Kolb described learners in terms of their paired preferences across both continua:

- accommodators: Phase 1 preference;
- divergers: Phase 2 preference;
- assimilators: Phase 3 preference;
- convergers: Phase 4 preference.

Please see print copy for Figure 3.2

Figure 3.2: Kolb’s Learning Cycle

In order to foster *critical thinking skills* (a focus in Chapters 6 and 7), Nieweg (2000) expanded Kolb's learning cycle. He proposed offering alternative sets of interpretations as inputs at the 'reflective observation' stage of the cycle. This was done to allow students support in such reflection. Their selections could be based on experience and prior knowledge. At the 'abstract conceptualisation' stage several explanations were also offered and students were asked at this stage to relate their selection to theory. Conclusions were then tested by returning to 'active experimentation'.

3.1.6 Finding a paradigm: an eclectic compromise?

It is the teacher's/researcher's belief that students construct knowledge in the dynamic environment in which the learner is situated. Impacting on the learning are the educational, physical, social, and cultural dimensions of the learner and those around him/her. Learning involves the acquisition of new knowledge and skill and the re-organisation of prior learning to accommodate the new knowledge. However, experience has also taught the teacher that the learning is not always that designed or intended by the teacher! Desired learning cannot be guaranteed, but through orchestration of effective paths, intent and practice, it can be effectively aligned to lead student learning to achievement of the required outcomes.

The researcher's classroom experience has allowed extensive observation of *what works* in facilitating student learning. The following list is by no means exhaustive, but is pertinent to this study:

- Students need a clear vision of *what they need t*o *know,* that is, *defined goals*;
- Students prefer to be actively engaged in learning and *experiential learning* promotes *deeper learning* through engagement in action and reflective consideration;
- Collaboration with teachers and peers accelerates student learning processes through the dynamic presentation of multiple perspectives. It also enhances confidence and intrinsic motivation;

- Structure can be provided to promote accommodation of new knowledge by facilitating formation of associations with the learner's prior learning/experience;
- Students may need to experience structure in their learning and this can be effectively achieved through sequencing, *scaffolding* or modelling;
- Independent practice affirms student learning by highlighting strengths and weaknesses in need of remediation or review;
- Reflection upon learning fosters independence, organisation of knowledge and the development of the higher order knowledge and skills associated with evaluation and judgement. Reflective practices also affirm the transfer of the skills that enable *lifelong learning*;
- 'Learning to learn' (*meta-learning*) promotes transferability of knowledge, facilitates the student's ability to learn discipline knowledge and skills and develops independence in the learner.

The teaching/learning framework adopted for this study has been constructed to accommodate *what works* and the discussion here may serve to situate the researcher's understanding of learning in the context of current learning theories. Table 3.1 presents a summary of the situation of such a framework in relation to these theories.

Table 3.1: Situating the researcher's teaching /learning framework

	Theory				
Provision for:	**Behaviourism**	**Cognitivism**	**Constructivism**	**Social constructivism**	**Co-constructivism**
Defined goals	√	√	√	√	√
Student led tasks	X[1]	√	√	√	√
Collaboration	X[1]	√[1]	√	√	√
Unstructured learning	X[1]	√	√	√	√
Structured learning	√	√	√	√	√
Independent practice	X[1]	√	√	√	√
Reflection	X[1]	√	√	√	√
Intrinsic motivation	X[2]	√	√	√	√
Feedback	√	√	√	√	√
Use of technology	√	√	√	√	√
Meta-learning	X[1]	X[1]	√[1]	√[1]	√

[1] Although not features of instruction under these theories, neither are they incompatible.

[2] Teacher instigated reinforcement, usually extrinsic in nature

Whilst the researcher values the social constructivist learning theories (and indeed the co-constructivist philosophy in particular), experience also supports merit in the work of both behaviourist and early cognitive theorists. The developmental history of learning theory has been mapped out by successful practice (or even unsuccessful practice highlighting limits in applicability) in the classroom and hence all of the above theories can afford construction of useful instructive frameworks. The learning framework selected for this project reflects an eclectic blend of successful strategies emanating from review of the literature and professional practice (discussed in further detail in Chapter 4), and supported by multiple aspects of the learning theories discussed in this chapter. A potential framework supported by the detail in Table 3.1 is illustrated in Figure 3.3.

Please see print copy for Figure 3.3

Figure 3.3: Teaching/learning framework showing relevant learning theory inputs.

(Adapted from Morris, Porter and Griffiths, 2004b, p.176)

3.2 Learning taxonomies

A question then arises: how is the desired learning to be detected? One approach is afforded by clear specification of learning outcomes in terms of observable behaviours which provide hallmarks of learning achievement. These behaviours do not represent the learning *per se* but are rather manifestations of that learning and hence are evidence of its taking place. Assessment then can record achievement of learning outcomes through recognition of these behaviours. However to re-draft an old adage: assessment only assesses what assessment assesses. Learning taxonomies, however, can provide a mechanism for evaluating the learning teachers have addressed in their teaching.

In judging the extent of student learning, teachers frequently ask: Do they know it? Do they understand? Have they grasped the key concepts? In turn it is relevant to reflect on whether all these questions asking the same thing or are the subtle differences important in reflecting upon the learning of students? This gives rise to the problem: How then is student learning to be described?

Initially there seemed little reason to focus on any assessment of motor or affective aspects of learning. Throughout the study, however, there was a growing realisation that involvement of the affective domain at university enhances development of *deep learning* by challenging individual interests and attitudes and grounding the learning in an environment supportive of community values.

Several taxonomies were considered for this project. Because of the discipline demands of mathematics and statistics (the teacher's primary arenas of practice), the researcher sought descriptions which incorporated a structural hierarchy of both knowledge and skills. Selection of an appropriate taxonomy was made after due consideration of two frequently used taxonomies: Bloom's Taxonomy and the SOLO Taxonomy.

3.2.1 Bloom's Taxonomy

Bloom's Taxonomy was developed to facilitate discussion of evaluation of student learning by providing:

> … a theoretical framework which could be used to facilitate communication among examiners. (Bloom, 1974, p. 4)

It aimed to achieve this by offering commonality of description. The taxonomy has also proved useful in building curriculum by aiding the specification of its aims and objectives and highlighting representation of different orders of knowledge and skills it fosters and the devise of assessment that addresses the aims and objectives of the curriculum. (Bloom, 1974) In this study a further use of the taxonomy has been to analyse the learning paths taken by students under instruction, and hence to evaluate any mismatch between instructional intent and student performance.

Bloom's uni-dimensional classification (1974) was based on behaviourally framed learning objectives. These objectives described the observable behaviours resulting from the targeted learning. Three domains were identified by Bloom:

1. Cognitive:

> … includes objectives which deal with the recall or recognition of knowledge and development of intellectual abilities and skills. (Bloom, 1974, p. 7)

2. Affective:

> ... includes objectives which describe changes in interest, attitudes, and values, and the development of appreciations and individual adjustment. (Bloom, 1974, p. 7)

3. Manipulative or motor- skill development: this area was not considered of immediate importance and hence development of this part of the taxonomy was deferred. (Bloom, 1974)

In its initial phase of development, the target was to construct an internally consistent and hierarchical classification of

> … the varieties of behaviours represented in the educational objectives… ((Bloom, 1974, p. 16)

Table 3.2: Condensed Version of the Cognitive Domain of the Taxonomy of Educational Objectives

Knowledge of:	
1.10 Specifics	
	1.11 Terminology
	1.12 Specific Facts
1.20 Ways and Means of Dealing with Specifics	
	1.21 Conventions
	1.22 Trends and Sequences
	1.23 Classifications and Categories
	1.24 Criteria
	1.25 Methodology
1.30 Universals and Abstractions in a Field	
	1.31 Principles and Generalisations
	1.32 Theories and Structures
Intellectual Abilities and Skills:	
2.00 Comprehension	
	2.10 Translation
	2.20 Interpretation
	2.30 Extrapolation
3.00 Application	
4.00 Analysis	
	4.10 Elements
	4.20 Relationships
	4.30 Organisational Principles
5.00 Synthesis	
	5.10 Production of a Unique Communication
	5.20 Production of a Plan or Proposed Set of Operations
	5.30 Derivation of a Set of Abstract Relations
6.00 Evaluation	
	6.10 Judgements in Terms of Internal Relevance
	6.20 Judgements in Terms of External Criteria

Adapted from Bloom, 1974, pp. 186-193

An abridged version of the Cognitive Domain of the Taxonomy is given in Table 3.2. It is divided into two sections: Knowledge; and Intellectual Abilities and Skills. Bloom has provided examples of related educational objectives for each of the sub-categories (Bloom, 1974).

Bloom's taxonomy was revised in 2001 (Anderson and Krathwohl, 2001). The revision cross-classified knowledge and cognitive processing skills and incorporated *meta-cognitive* knowledge. This inclusion was

> … predicated on our belief that it is extremely important in understanding and facilitating learning, a belief that is consistent with the basic precepts of cognitive psychology and supported by empirical research… (Anderson and Krathwohl, 2001, p. 44)

A condensed version of the Revised Taxonomy is given in Table 3.3. Although there has been some apparent loss of hierarchy between the categories, it is nevertheless evident within the sub-classifications shown in Table 3.4. In Appendix 3.1a and Appendix 3.1b, the teacher/researcher has split the sub classifications broadly into two groups representing 'lower order' and 'higher order' knowledge types. The sole basis for this 'split' has been the teacher's experience in high school and university teaching. However some support for this approach has been found in the work of delMas (2002) (see paragraph 3.4.2). Anderson and Krathwohl (2001) provided exemplars of the use of the revised taxonomy.

Table 3.3: Two-Dimensional Cross-Classification of Types of Knowledge by Cognitive Processing Skill

Please see print copy for Table 3.3

(Anderson and Krathwohl, 2001)

Table 3.4: Two-Dimensional Sub-cross-classification of *Conceptual* Knowledge by the Cognitive Processing Skill: *Understand*

Please see print copy for Table 3.4

(Anderson and Krathwohl, 2001)

Assessment can be designed to be representative of levels of knowledge and skills using Bloom's (or the revised) Taxonomy, but the consequent recognition of student learning may then be restricted to the pre-specified knowledge and skills. This limited applicability has given rise to much criticism (Biggs and Collis, 1979). However it must be remembered that the limitation lies in its lack of recognition of learning beyond the pre-determined assessment domain. The assessment need not, however, constrain student responses beyond that domain. Such *augmented* but *tangential* learning might still be recognised by comment/award or through imposition of additional criteria that allow for its recognition.

The aims of the assessment regimen need to be considered carefully before selection of an appropriate taxonomy. When the university and discipline aims can be specified behaviourally, then the revised taxonomy can prove effective in designing assessment and indeed in aligning instruction with the desired learning represented in that assessment. However there may be need for a more broad-based assessment of student learning which allows qualitative evaluation of student responses.

3.2.2 SOLO Taxonomy

While noting that Bloom's Taxonomy was extremely useful for devising assessment that covered all aspects of knowledge and cognitive processing, Biggs and Collis expressed concern that this very function restricted recognition of student performance under assessment to those behaviours specified *a priori*. They proposed taxonomy based upon the observed behaviours evident in assessment responses, the SOLO (*s*tructure of the *o*bserved *l*earning *o*utcome) Taxonomy. Their learning taxonomy was based on the *quality* of student responses, and hence was applied ad hoc to student responses on a given task at a given time (Biggs and Collis, 1982, Biggs and Collis, 1979).

The SOLO Taxonomy has its foundations in developmental psychology. Biggs and Collis identified an isomorphism between the taxonomy and the developmental stages of Piagetian Theory. Table 3.5 illustrates this relationship. The stages (and the response descriptions) are strictly hierarchical, each stage presuming achievement at the previous one.

Table 3.5: Isomorphic mapping of Piaget's Stages of Cognitive Development to the SOLO descriptions

Developmental Stage with Minimal Age	SOLO Description	Relation Operation	Consistency and Closure
Formal Operations (16+ years)	Extended Abstract	Deduction & induction; Generalises beyond experience.	Inconsistencies resolved Conclusions left open for possible alternatives
Concrete Generalisation (13-15 years)	Relational	Induction; Generalise within experience using related aspects.	No inconsistency within the closed system
Middle Concrete (10-12 years)	Multi-structural	Generalises in terms of limited, independent aspects.	Awareness of consistency, but potentially alternative conclusions
Early Concrete (7-9 years)	Uni-structural	Generalises in terms of one aspect	No need for inconsistency; jumps to conclusions
Pre-operational (4-6 years)	Pre-structural	Denial; tautology; transduction; bound to specifics.	No need for consistency; failure to identify the problem.

(Adapted from Biggs and Collis, 1979)

The SOLO taxonomy may have been grounded in developmental psychology, but in its application to response evaluation, Biggs and Collis also found wide variation in student performances across stages (both for group responses on the same task and individual responses across various tasks). They proposed two domains: the internalised and developmental aspects of learning peculiar to the individual – the *hypothetical cognitive structure* and the overt response to the learning tasks – the *structure of observed learning outcomes* (SOLO) peculiar to the task. (Biggs and Collis, 1982, Biggs and Collis, 1979)

It is evident that the taxonomy offers pertinent criteria for essay type responses or those where the cognitive approach is overt; however application in the discipline of mathematics/statistics appears more problematic. For example, the authors' discussion of this application in algebra disclosed a dilemma:

> "You are to decide whether the following statements are true always, sometimes or never. If you put a circle around 'sometimes' explain when the statement is true" All letters stand for whole numbers or zero (e.g. 0, 1, 2, 3 etc.)
>
> | | |
> |---|---|
> | 1. a + b = b + a | Always |
> | | Never |
> | | Sometimes, that is when____ |
> | 2. m + n + q = m + p + q | Always |
> | | Never |
> | | Sometimes, that is when____ |
> | 3. a + 2b + 2c = a + 2b + 4c | Always |
> | | Never |
> | | Sometimes, that is when____ |
>
> (Biggs and Collis, 1979, p. 91)

As can be seen in Table 3.6, Uni-structural, Multi-structural and Relational responses for question 1 were virtually indistinguishable. Indeed the Uni-structural and Multi-structural responses were identical across the three questions making determination of performance level impossible. Disclosure of the rationalisation processes was needed in order to discriminate between the response levels. This type of explanatory response is rarely available in current assessment regimens in the tertiary sector. Thus in order to achieve the desired *discriminability*, there needs to be a great deal of forethought in task construction.

Table 3.6: Classification of student responses to the above algebra questions using the SOLO Taxonomy.

Question	Responses			
	Uni-structural	Multi-structural	Relational	Extended abstract
1	Always Students tried only one number	Always Students tried several numbers	Always Students regarded the pronumeral as generalised, but no notion of a variable	Always Notion of a variable was evident
2	Impossible for student at this level	Impossible for student at this level	Never Unable to accept that n=p	Sometimes When n=p
3	Impossible for student at this level	Impossible for student at this level	Never 0 did not seem a possibility	Sometimes When c=0

Adapted from (Biggs and Collis, 1979, p. 91-94)

Analysis of the responses can be tedious and time consuming even when assessment has been carefully constructed to afford opportunities for student display of the necessary diagnostic features. The questions then might be asked: is the SOLO Taxonomy not then susceptible to the criticisms aimed at Bloom's Taxonomy? Is not the learning behaviour then to be targeted *a priori*?

Although the SOLO Taxonomy allows classification of learning in terms of the *quality* of student cognitive construction of their responses, it appears to afford no consideration of the hierarchy of complexity in the structure of the underlying knowledge presented in the instructional processes.

The two dimensional perspective offered by the revised Bloom's Taxonomy selected for this research, identifies the behaviours resulting from student *cognitive processing* (from simple to complex) of *types of knowledge* (also from simple to complex) evident in predetermined and specified behavioural objectives for assessment (Anderson et al., 2001). Although each of the two taxonomies under consideration display merit in the recognition of student learning, each also presents difficulty in practical application in the disciplines of mathematics and statistics.

3.2.3 A contextualised approach to classification?

The most difficult aspect of applying either of the considered taxonomies proved to be maintaining consistency. Language can be rich in meaning when individuals bring their personal connotations to the interpretation of basic tasks in assessment. Bloom's Taxonomy may be used to classify objectives for assessment by categorising the *language of the expected behaviours* according to those given in the Handbook. However the researcher and her colleagues still encountered discrepancies in individual classifications:

> … Simple requests such as 'explain' contain so many subtle nuances dependent upon the context. A student may be asked to 'explain using the output given'. This requires discriminating between relevant and irrelevant information and using it to *justify*. Alternatively, the students may be asked to 'explain' a concept with the aid of a diagram (*interpret*), or to 'explain' a process i.e. list the steps involved (*recall* a procedure). These all demand different levels of skills. It would appear that without discipline specific vocabulary, such a general taxonomy is bound to prove difficult to use and use reliably! (Morris, Porter, and Griffiths, 2004a, p.15)

There is scope for development of discipline specific (indeed possibly even subject specific taxonomies). Consideration could be given to using statistical analyses to generate groupings of student assessment responses based upon most frequently correct to least. Deconstruction of the language of the questions generating the groupings may provide commonalities (or else highlight misinterpretations) which could be used to structure a taxonomy. Such a task is however beyond the field of this study.

3.3 Learning styles

The term *learning styles* has been coined by many researchers differently. Osborn and Plunkett (2003) refer to learner preferences for selected styles of assessment. Other researchers refer to preferred cognitive styles (Riding and Rayner, 1999) classified across two dimensions: wholist-analytic and verbal-imagery. Mainemelis, Boyatzis and Kolb (2002) refer to learning styles which align closely with the experiential learning theory of Kolb and proponents describe learners as activists, reflectors, theorists and pragmatists (Watkins et al., 2002). Because of the researcher's support of the precepts of experiential learning theory, this is the definition of learning styles that will be referred to in this work.

Mainemelis et al., (2002) tested differences in adaptive flexibility between balanced and specialised learning styles. Their results detected greater adaptive flexibility in the balanced group for the *perception* dimension (experiencing and conceptualising contexts) and the

> … learning style specializing in experiencing showed higher levels of skill development in interpersonal skills and lower levels of skill development in analytic skills; while the reverse was true for the learning style specializing in conceptualizing. (Mainemelis, et al., 2002, p. 5)

but they found no consistent differences for the *processing* dimension (acting and reflecting contexts). (Mainemelis, et al., 2002)

Learners have also been described in terms of their preferred mode of reception such as visual or auditor. However as Watkins et al. (2002) claim, to define learners in such facile and uni-dimensional terms may be limiting and a rich learning environment which affords exposure to multiple styles allows development of all facets of the learner.

3.4 Statistical learning

The teacher/researcher's primary fields of teaching expertise are mathematics and statistics. One of the case studies for this research involved teaching introductory statistics in undergraduate degree programs. It was thus expedient to investigate learning goals for this field. Cobb and Moore (1997) claim

> … we may focus on content, what we want our students to learn, or on pedagogy, what we can do to help them learn. These two topics are of course related. In

> particular, changes in pedagogy are often driven in part by changing priorities for what kinds of things we want our students to learn. (Cobb and Moore, 1997, p. 11)

Content is discussed briefly in this section and pedagogy will be examined in the next chapter.

3.4.1 Defining the learning

Although many novices approach statistics as *yet another maths subject*, there are some sharp contrasts between the two disciplines (Moore and Cobb, 2000; Cobb and Moore, 1997; Gal and Garfield, 1997). Many researchers have highlighted the importance of context in statistics:

> Statistics requires a different way of thinking, because data are not just numbers, they are numbers with a context. (Cobb and Moore, 1997, p. 801)

Context assumes a totally different role in mathematics on the other hand.

> ... the ultimate focus in mathematical thinking is on abstract patterns: the context is part of the irrelevant detail that must be boiled off over the flame of abstraction in order to reveal the previously hidden crystal of pure structure. In mathematics context obscures structure… In data analysis, context provides meaning. (Cobb and Moore, 1997, pp. 802-803)

Another source of confusion for students enrolled for the first time in statistics is that mathematics frequently offers a single solution to a problem. Statistics on the other hand may offer multiple reasonable alternatives. (Gal and Garfield, 1997)

Moore and Cobb (2000) claimed that

> … statistics is a subject whose goal is to solve real-world problems. (Moore and Cobb, 2000, p. 617)

Consequently, statisticians find employment in diverse workplaces. In academia too, their research is affiliated with numerous disciplines. Perhaps this exposure to disciplines other than their own has encouraged statistics educators to be innovative in reforming teaching practice, particularly in elementary and service courses (Gal and Garfield, 1997) while such practices have been slower to take hold in mathematics. (Moore and Cobb, 2000; Cobb and Moore, 1997)

Learning goals framed in constructivist philosophies have percolated through from education faculties to science and mathematics. The focus in statistics has also shifted to include

> … "more data and concepts, less theory, fewer recipes" ... more active learning opportunities and … decrease the amount of lecturing. (Gal and Garfield, 1997, p. 2)

If the desired learning outcome were *statistical thinking* or emulation of the work of practising statisticians, then such strategies might conceivably have been perceived as more useful (Rossman, 1997).

Bellhouse (2003) advised the teaching of the history of statistics because

> … in using history in probability and statistics the important question to address is: how and why was this new knowledge created? (Bellhouse, 2003, p. 1)

He believed that

> … the best use of history in class is to discover and describe what motivated people to work on different problems. And it will often turn out that what motivated our statistical forbears to come up with certain techniques or theory is the same as the motivation for using the results and techniques today. (Bellhouse, 2003, p. 7)

Chance (2002) distinguished between *statistical reasoning*, *statistical literacy* and *statistical thinking* which she regarded as encompassing the others. She canvassed several definitions, narrowly defining *statistical literacy*

> … as understanding and interpreting statistical information presented, for example in the media … (Chance, 2002, p. 4)

and *statistical reasoning*

> … as working through the tools and concepts in the course … (Chance, 2002, p. 4)

and concluded that *statistical thinking* includes *what a statistician does* and involves

> … (moving) beyond constructing a plot, solving a particular problem, reasoning through a procedure and explaining a conclusion. Perhaps what is unique to statistical thinking, beyond reasoning and literacy, is the ability to see the process a whole (with iteration) including the "why", to understand the relationship and meaning of variation in this process, to have the ability to explore data in ways beyond what has been prescribed in texts, and to generate new questions beyond those asked by the principal investigator. (Chance, 2000, p. 6)

Rumsey (2002) has claimed that definitions of *statistical literacy* have been too broad and instead referred to two distinct learning outcomes:

> "Statistical competence" refers to the basic knowledge that underlies statistical reasoning and thinking, and "statistical citizenship" refers to the ultimate goal of developing the ability to function as an educated person in today's age of information. (Rumsey, 2002, pp. 3-4)

Her second outcome might be perceived as encompassing both *statistical reasoning* and *statistical thinking* (see figure 3.4).

Teaching *statistical thinking* has therefore necessitated movement away from *telling* students about it, to engaging them in a more experiential learning environment. In elementary subjects, students of statistics have been increasingly exposed to *experimentation with data* as a means of devolving the processes of *statistical thinking*. The discovery process is used to develop understanding and appreciation of the tools and procedures of statistics. (Moore and Cobb, 2000; Chance, 2000) Rossman and Chance (1999) have described their recommendations for incorporating statistical inference with the teaching of exploratory data analysis and data collection. They claimed that their

> … goal is to focus on … student investigation and discovery of inferential reasoning, proper interpretation and cautious use of results, and effective communication of findings. (Rossman and Chance, 1999, p. 298)

In recent times, there has been less emphasis on the mathematical origins, and mechanical and arithmetic applications of the theoretical bases. The drudgery of calculation has been relieved by the use of technology, freeing students to focus on the thinking processes behind the actions. (delMas, 2002; Chance, 2002; Moore and Cobb, 2000)

Gal and Garfield (1997) identified the knowledge and skills desired as outcomes of elementary courses in statistics. They recommend the development of students' statistical knowledge and skills so that they can:

> - Comprehend and deal with uncertainty, variability, and statistical information in the world around them, and participate effectively in an information-laden society.
> - Contribute to or take part in the production, interpretation, and communication of data pertaining to problems they encounter in their professional life. (Gal and Garfield, 1997, p. 3)

The pedagogical structures that support these goals have been discussed in Chapter 4.

3.4.2 Classifying the learning

Rumsey (2002) identified two broad goals of introductory statistical instruction: ‘statistical citizenship’ and the development of scientific research skills. She perceived

> … good “statistical citizens” (as) understanding statistics well enough to be able to consume the information they are inundated with on a daily basis, (thinking) critically about it, and (making) good decisions based on that information. (Rumsey, 2002, p.1)

Achievement of these goals, however, requires more than the acquisition of the basic knowledge of ideas, terms and language, but also includes the more complex cognitive demands of evaluation, decision making, judgement and explanation. Hence achievement of the goals necessitates aspects of reasoning and thinking. Rumsey reinforced this idea, claiming that in order for individuals to be effectively engaged

> … in the chain of statistical information, there will be a need for a basic understanding of the concepts and language, a level of reasoning (the abilities to question, compare and explain) and a level of statistical thinking (applying the ideas to new problems and identifying questions of your own). (Rumsey, 2002, p. 2)

Her descriptions of *statistical reasoning* and *thinking* align with those of Chance (2002), delMas (2002) and Garfield (2002). Rumsey’s (2002) definition of *statistical literacy* included statistical ‘competence’ and ‘citizenship’ The inclusion of ‘citizenship’ has extended the notion of *statistical literacy* from basic competence into more complex cognition. Although Rumsey has attempted to describe basic competence this has been achieved largely through discussion of examples and no attempt was made to specify any hierarchical structure in the knowledge and skills involved.

Garfield (2002) has focused on the processes of *statistical reasoning*, also noting that the term has been used interchangeably with *statistical thinking*. Her model of *statistical reasoning* has been included in Appendix 3.2. It contains an inherent hierarchy, which describes five levels of *statistical reasoning*. She elaborated on the model to include implications for instruction and assessment.

delMas (2002) compared and contrasted the discussions of the domains of the cognitive outcomes of statistical learning proposed by Rumsey, Garfield and Chance. He illustrated two models recognising overlap between pairs of the domains using Venn diagrams and these have been illustrated in Figure 3.4. The first diagram highlights some overlap of three independent domains. The second diagram represents an overarching domain of *statistical literacy* that includes the other two domains. This diagram is commensurate with Rumsey's definition of *statistical literacy*. He also postulated other possibilities representing *statistical reasoning* or *statistical thinking* as the dominant domain encompassing the others. Of course, other representations also spring to mind. Perhaps it is the objectives of each subject that will dictate the relative importance of each of the cognitive domains.

Please see print copy for Figure 3.4

Figure 3.4 Outcomes of statistics education

Source: delMas, 2002, p.4

In addition, delMas (2002) has claimed that

> What moves us from one of the three domains to another is not so much the content, but, rather, what we ask students to do with the content. (delMas, 2002, p. 5)

Tasks that promote development of *statistical literacy* ask

> … students to identify examples or instances of a term or concept, describe graphs, distributions and relationships, to rephrase or translate statistical findings, or to interpret the results of a statistical procedure. (delMas, 2002, p. 5)

In encouraging students to

> … explain how or why results were produced … or why a conclusion is justified … (delMas, 2002, p.5)

it is *statistical reasoning* that is targeted. However, *statistical thinking*

> … asks students to apply their basic literacy and reasoning in context … (and)… is promoted when instruction challenges students to apply their understanding to real world problems, to critique and evaluate design and conclusions of studies or generalize knowledge obtained from classroom examples to new and somewhat novel situations. (delMas, 2002, pp. 5-6)

These behaviours have been summarised in Table 3.7. The researcher has noted the order of cognitive processing skills identified by delMas and has identified the similarities to her own 'splitting' of the cognitive processing skills of the revised taxonomy of Bloom (Anderson and Krathwohl, 2001) as shown in Appendices 3.1a and 3.1b. The skills associated with development of *statistical literacy* generally correspond with her identified 'lower order skills' whilst those associated with *reasoning* and *thinking* also broadly correspond with her 'higher order skills'.

Table 3.7: Classifying statistical learning: the three domains

Please see print copy for Table 3.7

Source: delMas, 2002 p.6

The researcher is in accord with the work of Chance, Garfield, delMas, Rumsey and others in their search for a taxonomy that affords more effective classification of statistical learning, there is more work to be accomplished before this can be achieved, and as the revised taxonomy is not at odds with current statistical learning research, it has been selected to provide the mechanism for defining statistical learning in this study.

3.5 From learning to instruction

Armed with an understanding of what learning is and, in particular, a working definition of the targeted learning in this project, the researcher sought to construct a framework focused on promoting that learning. The next stage of her research involved a review of the literature canvassing potential strategies valued for their success in achieving these specific learning outcomes.

Chapter 4

Teaching for learning: as easy as falling off a horse!

Whenever the horse stopped (which it did very often), [the White Knight] he fell off in front; and, whenever it went on again (which it generally did rather suddenly), he fell off behind. Otherwise he kept on pretty well, except that he had a habit of now and then falling off sideways; and as he generally did this on the side on which Alice was walking, she soon found that it was the best plan to not to walk *quite* close to the horse.

"I'm afraid you've not had much practice at riding," she ventured to say, as she was helping him from his fifth tumble.

The Knight looked very much surprised, and a little offended at the remark. "What makes you say that?" he asked, as he scrambled back into the saddle, keeping hold of Alice's hair with one hand, to save himself from falling over on the other side.

"Because people don't fall off quite so often, when they've had much practice."

"I've had plenty of practice," the Knight said very gravely: "plenty of practice!"

Alice could think of nothing better to say than "Indeed?" but she said it as heartily as she could. …

"The art of riding," the Knight suddenly began in a loud voice, waving his right arm as he spoke, "is to keep –" Here the sentence ended as suddenly as it had begun, as the Knight fell heavily on the top of his head exactly in the path where Alice was walking. She was frightened this time, and said in an anxious tone, as she picked him up, "I hope no bones are broken?"

"None to speak of," the Knight said, as if he didn't mind breaking two or three of them. "The great art of riding, as I was saying, is – to keep your balance properly. Like this, you know – "

He let go of the bridle, and stretched out both his arms to show Alice what he meant, and this time he fell flat on his back, right under the horse's feet.

"Plenty of practice!" he went on repeating, all the time that Alice was getting him on his feet again. "Plenty of practice!"

"It's too ridiculous! cried Alice, losing all her patience this time. "You ought to have a wooden horse on wheels, that you ought!"

From "It's my own invention" in
Through the Looking Glass by Lewis Carroll, p. 208

4.0 Instructing to construct desired learning

The central role of the teacher is to engage the student in a journey that will culminate in achievement of the specified learning goals. These goals are usually formulated under pressure from various interested groups: discipline, university, community, professional etc. Whilst universities have defined common goals under such pressures, the development of accompanying pedagogical structures has not always resulted in commonality of approach. Several key concepts in pedagogical design underpin this study:

- *Alignment* of teaching, learning and assessment with defined learning goals;
- Active learning engaging students in higher order thinking and this in turn promotes *deeper learning*;
- Defined and achievable learning goals provide students with focus and motivation and
- Promotion of *meta-cognitive* skills to enabling life-long learning.

4.1 Defining concepts

Throughout this study, italics have been used to highlight particular concepts related to teaching, learning and assessment. The literature may have evidenced variation in definition or uses of these terms and hence the researcher has included an explanation of their explicit meaning within the current study.

4.1.1 *Meta-cognition*

Hattie, Biggs and Purdie (1996) identify *meta-cognition* as that cognitive faculty associated with

> … self management of learning, that is, … planning, implementing, and monitoring one's learning efforts, and … the conditional knowledge of when, where, why, and how to use particular tactics and strategies in their appropriate contexts. (Hattie, Biggs and Purdie, 1996, p. 99)

Anderson and Krathwohl (2001) identify two separate dimensions in the notion of *meta-cognition*:

> … knowledge about cognition and control, monitoring and regulation of cognitive processes. (Anderson and Krathwohl, 2001, p. 43)

Since the latter refers to processing, it has been absorbed by the cognitive processing dimension of their taxonomy which has been used in this study. Only the former is addressed by the authors as *meta-cognition.*

The revision of Bloom's Taxonomy by Anderson and Krathwohl (2001) included the new knowledge category of *meta-cognition.* The authors claimed that this inclusion has followed from the prevailing interest in students' knowledge about their own learning processes arising from cognitive and social constructivist theories. Many researchers have claimed that *deeper learning* results from opportunities for students to reflect upon their own learning: assessing strengths, weaknesses and motivation. In a climate of rapidly increasing availability of new ideas and practices and the need for professionals to constantly update their own knowledge bases, developing appropriate learning strategies promotes a facility for lifelong learning.

Table 4.1: Two-Dimensional Sub-classification of *Meta-cognitive* knowledge

Meta-cognitive knowledge	Strategic knowledge: knowledge of the general strategies for learning, thinking and problem solving.
	Knowledge about cognitive tasks: knowledge of the complexity of cognitive demands of different tasks
	Self-knowledge: assessment of personal strengths and weaknesses

(Adapted from Anderson and Krathwohl, 2001, pp 55-60)

The lack of focus upon the Affective domain of Bloom's taxonomy has caused concern among educators, particularly those wishing to inculcate 'appreciation' and 'values' in learners. Anderson and Krathwohl (2001) believe that their *meta-cognitive* category has provided, in some measure, a bridge to the Affective domain that might be developed by other researchers in future revisions.

4.1.2 *Deeper learning*

Biggs (1999) tracked usage of the term *deeper learning* to the study of Marton and Säljö published in 1976. Their observations identified two different approaches of students exposed to text for which they anticipated questions. The first group aggregated disjointed facts and details on which they believed they might be quizzed. These learners read superficially and could not respond to questions beyond content. The second group, however, retrieved the meaning and purpose of the author. They searched 'beneath' the surface of the text for *deeper* meaning. These students exhibited *deeper learning* while the others only demonstrated *surface learning.* (Biggs, 1999, Marton and Säljö, 1976)

Clearly *deeper learning* engages higher order thinking skills involving analysis and critical evaluation, skills much prized as learning outcomes at the university level. However as Biggs has claimed

> … meaning is not imposed, or transmitted by direct instruction, but is created by the student's *learning activities*, well summarised in the term "approaches to learning". A surface approach refers to activities of an inappropriately low cognitive level, which yields fragmented outcomes that do not convey the meaning of the encounter. The deep approach refers to activities that are appropriate to handling the task so that an appropriate outcome is achieved. The surface approach is therefore to be discouraged, the deep approach encouraged- and that is my working definition of good teaching. (Biggs, 1999, p.60)

A *deeper* approach encourages transformative learning – learning that develops the individual through encounters with multiple perspectives and results in personal awareness and a realisation of the individual's role in community. (Cranton and Roy, 2003) Through active engagement in such learning, the learner's conceptual perceptions of the world are altered.

> Thus education is about conceptual change, not just the acquisition of information. (Biggs, 1999, p. 60)

Biggs added that conceptual change can only occur when

> It is clear to students (and teachers) what is "appropriate", what the objectives are, where all can see where they are supposed to be going, and where these objectives are buried in the assessment tasks. (Biggs, 1999, p. 60)

He also cautioned that students need to be motivated, task-focused not test-focused and able to work collaboratively with both teachers and peers.

The concept of *deeper learning* is discussed more fully in Chapter 7, since in that case study *surface learning* was a primary concern to the teachers involved. This brief discussion has been presented here as the teacher has deemed *deeper learning* to be of primary concern to any "good teaching", and hence some clarity of definition is of import to the entire study.

4.1.3 *Organisers*

The key concepts underpinning Ausubel's learning theory is the organisation and retention of meaningful new knowledge. His experiments (Ausubel, 1960) supported the introduction of related generalised concepts prior to the introduction of new knowledge to students. He explained that this increased the effectiveness of learning by facilitating the student's ability to incorporate the knowledge and retain it. He claimed that this is achieved indirectly

> …by providing 'ideational scaffolding (Ausubel, 1978, p. 253)

The *organiser* provides a generalised structure that permits greater interpretability of the new material and grounding for its more permanent retention. They

> … consist of introductory material at a higher level of abstraction, generality, and inclusiveness than the task itself. . The function of the organizer is to provide ideational scaffolding for the stable incorporation and retention of the more detailed and differentiated material that follows ... as well as to increase discriminability between the latter and related, interfering concepts in cognitive structure. (Ausubel, 1968, p. 29)

Ausubel conceived *advance organisers* to be successful in:

> (a) the selective mobilization of the most relevant existing concepts in the learner's cognitive structure for integrative use as part of the subsuming focus for the new learning task, thereby increasing the task's familiarity and meaningfulness; and (b) the provision of optimal anchorage for the learning material in the form of relevant and appropriate subsuming concepts at a proximate level of inclusiveness. (Ausubel, 1960, p. 271)

An example of an *advance organiser,* applicable in the case study reported in Chapter Six in this work, is illustrated in Table 4.2. It has been designed by the teacher to be given to students in a fundamental statistics class and introduces students to the *process* of exploring the relationship between two variables. The content of the organiser has been selected for its

> … suitability for explaining, integrating, and interrelating the material …" it precedes (Ausubel, 1968, p.81).

The illustrated generalised structure would be equally applicable (and hence indicative of a higher level of abstraction) for relationships either between numeric variables or between categorical variables and hence this *organiser* would be used prior to discussions/lectures in both cases. In this way it also identifies associations between the two sets of processes. It would also be used in reviews as a prompt. The italicised words indicate processes already experienced in the unit of study and hence are already familiar to the learner. A more complete description of the explorative process for numeric variables is elaborated in Table 4.3.

Table 4.2: Example of an *organiser* for exploring the relationship between two variables

1. *Explore* the potential for a relationship
2. *Determine* the form of the relationship
3. *Evaluate* the significance of the relationship
4. *Discuss* the limitations of the relationship
5. *Conclude* based on the examined evidence

Anderson et al. (1978) maintained that Ausubel's investigations into the effect of *advance organisers* have been inconclusive. They claimed that an individual's existing schemata alone will determine meaning in text and an obscure set of abstract and unfamiliar terms would do little to promote interpretability of new knowledge. However, they did not present the potential of an *advance organiser* framed in familiar concepts or processes as highlighted by Ausubel himself. He prescribed that the concepts be

> … relatable to presumed ideational content in the learner's current cognitive structure. (Ausubel, 1978, p.252)

Ausubel also reminded his critics that the *organisers* are not useful for rote learning of disconnected material, but rather for 'meaningful learning' of more complex but connected material. (Ausubel, 1978)

4.1.4 *Scaffolding*

Whilst Ausubel's *organiser* offers 'ideational scaffolding' from a higher level of abstraction than the instructional detail, *scaffolding* is generally regarded as supporting concept development throughout the instructive period. *Scaffolding* is inclusive of Ausubel's notion of the *organiser*, but not restricted to it. *Modelling* may provide *scaffolding*. Alternatively, the detailed elaboration shown in Table 4.3 of the *organiser* given in Table 4.1 may support student learning of the required exploratory process. Note that the detail may be removed once students have reached an expected level of proficiency in performing complex tasks. It should also be noted here that this example of *scaffolding* is not meant to provide comprehensive support for this particular topic.

Further *scaffolds* may emanate from points within this structure to encompass more in depth treatment. For example, no detail has been presented for approaches to non-linearity involving transformations. Neither has residual analysis been included in 'Limitations of the model'. These considerations would require a further set of *scaffolded* steps. However, incorporating these 'extra' supports into the primary structure would present a potentially confounding complexity in the introductory process and could prove counterproductive as a support to learning.

As the *organiser* provides focus structures for the assimilation of new knowledge and skills, it could possibly be likened to a filing system. *Scaffolding* affords detailed support structures to facilitate the learning of complex knowledge and skills. This perhaps relates to the labelled folders that reside within the filing system.

Table 4.3: Example of *scaffolding* for exploration of a linear relationship between two numeric variables

1. ***Explore*** the potential for a relationship
   - Construct a scatter plot
     - Check for non-linearity
     - Comment on **strength and direction of the relationship**
       - From the scatter plot
       - Determine the value of the correlation coefficient
       - Explain the significance of the value of the correlation coefficient
       - In terms of its size
       - In terms of its sign
     - Check for outliers/influential points
     - Check for alternative models
     - Check for patterns
2. ***Determine*** the form of the relationship
   - Identify the dependent and independent variables
   - Find the **equation** of the line of best fit
     - Determine the value of the intercept
     - Determine the value of the slope
     - Write the equation for the relationship between the two variables
3. ***Evaluate*** the significance of the relationship
   - Determine the value of the coefficient of determination
   - Interpreting the value of the coefficient of determination
   - Examining the t tests on the parameter estimates
     - Constant
     - Slope
   - Analysing variation (ANOVA)
     - Explained by the model
     - Not explained by the model
     - The F statistic
     - The p value for the F test
4. ***Discuss*** the limitations of the relationship
   - The domain of values
   - The variation
   - The model structure
5. ***Conclude*** based on the examined evidence
   - In the **context** of the problem and including **evidence** from:
     - Results of '**exploration**'
     - Results of '**evaluation**'
     - While recognising the **limitations**
     - Stating the '**determined**' form of the relationship

Vygotsky's constructivist theory (Ussher and Gibbes, 2002) placed emphasis on the cooperative nature of classroom learning, with students' attainment dependent upon the interactions of teacher, peers, the environment and the learner themselves. His perceived role of the teacher was, however, a particular one: the mediator and guide of the learning process. This guided learning is achieved through *scaffolding* that supports the learner from their existing knowledge context through to the next level. (Daniels, 2003, Ussher and Gibbes, 2002, Wood and Wood, 1996)

Scaffolding should limit the frustration arising from confrontation with an overall learning goal presently beyond the learner's grasp. It maintains focus on the ultimate goal while effectively guiding the learner through the intermediate steps to its attainment. Appropriate structures gradually shift the responsibility from the teacher to the learner. (Wood and Wood, 1996)

4.1.5 *Alignment*

> *Alignment* refers to the degree of correspondence among the objectives, instruction and assessment. (Anderson,and Krathwohl, 2001, p. 10)

In order to foster student achievement of the specified learning outcomes, instruction should target the desired learning either directly or through provision of supportive scaffolding. Fair assessment should also target these outcomes. (Anderson and Krathwohl, 2001)

> Making the assessment process transparent, and aligning it with goals and learning activities, are critical steps in promoting improved learning outcomes. (Angelo T., 1999, p. 11)

Biggs (1999) recommended a teaching/learning environment founded upon the explicitly defined 'verbs' in the objectives. He too, cautioned that assessment needs to target achievement of these objectives.

> In making our objectives clear it is essential that we unpack and make explicit the meanings we want our students to address. (Biggs, 1999, p. 66)

Furthermore, these

> … objectives need to be stated in such a way as to allow the information from the assessments to specify the level of pass. (Biggs, 1999, p. 66)

Biggs claimed

> Constructive alignment is common sense, yet most university teaching is not aligned. (Biggs, 1999, p. 73)

and attributed this to the traditional expository teaching styles yet to be questioned as effective by many academics. His position would seem to have been strengthened by the broader perspective of others in higher education (Yorke, 2001; West, 1998; NCIHE, 1997) who, while acknowledging the academic expertise of teaching staff in their discipline, also recognised their limited training in teaching

> Academics are all committed to keeping abreast of the latest research and ideas in their discipline, but few of them have the opportunity to keep at the forefront of developments in how to teach their subject. … only just over half of academics have ever received any training in how to teach and over two thirds of those had received training only at the beginning of their careers. This inevitably means that a large proportion have had no training in, for example, the use of information technology for learning and teaching. (NCIHE, 1997, Section 3.4)

Alignment can be achieved through application of an appropriate learning taxonomy. Peremptory examination of all aspects of teaching, learning and assessment can be deceptive in determining the degree of alignment. The rigour afforded by application of a taxonomy of learning to the *intent* of teaching (evidenced in the objectives of teaching, the learning tasks and assessment questions) and the *practice* of teaching and learning (evidenced in the delivered lectures, the task and assessment solutions and any corresponding marking criteria) serves to identify any *cognitive mismatches* (discrepancies between the knowledge and skill level in the subject's presentation and that of the cognitive demand of learning tasks and assessment) (Lugenbehl, 2003; Biggs, 1999). Biggs (1999) claimed that

> Lack of alignment is a major reason why students adopt a surface approach to learning. (Biggs, 1999, p. 69)

He also maintained that even appropriately structured objectives do not prevent *misalignment* when results of the assessment are required to be reported quantitatively or 'spread along a curve'. (Biggs, 1998)

Misalignment will also be reflected in student performance in assessment (Lugenbehl, 2003). For example, if the outcomes require students to critically evaluate, then assessment should target this learning, but the instruction needs to model the expected behaviour and hence should not superficially report summaries of background material or litanies of historical fact without demonstrating critical reflection upon them. Johnstone (2005) specifically identified the *misalignment* that results from an expository model of teaching that promotes *shallow learning,* but still expects the critically reflective responses in assessment that typify *deeper learning*. These mismatches generate intense frustration in the learner (Johnstone, 2005; Lugenbehl, 2003).

4.2 Assessment

Piper et al. (1996) reported on an extensive survey of Australian university examiners about their examination practices. The survey results revealed a predominantly traditional approach to assessment, with little reference to competency standards or innovative practice, although the survey instrument made no provision for recognising them. Responses indicated that assessment has been driven more by tradition and expediency rather than the address of learning outcomes. Of even greater concern is that only 19% of the respondents recognised the need for changes to assessment practices! Deficiencies in training for examiners were also highlighted. There was little evidence of understanding of the development of appropriate marking criteria or their application. (Piper et al., 1996)

In order to understand current assessment practices, it is necessary to go beyond such strong evidence of traditional practices, and also examine the impact of other stake holders. Political, professional and community expectations have pressured the focus of assessment towards accountability. Despite such concerning issues, many enlightened educators still envisage improvement in student learning as a prime focus. (Brew, 1997)

Australian universities have been confronted by increased student intakes, no proportionate increases in staff funding, a greater representation of international students (West, 1998), professional and societal demands for addressing team skills and enhanced opportunities for plagiarism (James et al., 2002). James et al. (2002) have consequently identified five key assessment issues facing higher education:

1. Exploiting the potential of online assessment;
2. Design of efficient and effective assessment for large classes;
3. Fostering academic honesty;
4. Using assessment to guide group work;
5. Supporting novices to Australian university assessment practices.

(Adapted from James et al., 1996, pp. 4-7)

Underpinning current trends in improving assessment at universities is the need for teachers

> … first to be clear about what we want our students to learn, and then teach and assess accordingly in an *aligned* system of instruction. (Biggs, 1999, p. 64)

Angelo (1999), however, has cautioned that in the evaluation of educational programs,

> … our assumptions about learning outcomes should be empirically tested and that our claims should be based in evidence. (Angelo, 1999, p. 5)

Most students see assessment as defining the curriculum. This then positions it as of strategic value to student learning. (Biggs, 1999; James et al., 1996) Authentic assessment that targets critical thinking focuses students on developing *deeper learning*. Equally, assessment that seeks recall or repetition of fact merely encourages *surface learning* (Biggs, 1999; Marton and Säljö, 1976). Defined outcomes for assessment promote self-assessment; while clearly signposted marking criteria for assessment can afford detailed feedback to students on achievement (see paragraphs 4.3.6 - 4.3.7).

Despite the apparent gloom of the DEETYA report (Piper et al., 1996), the authors concluded with an extensive inventory of 'best practice' for assessment. Similarly

James et al. (1996) have described a set of indicators of effective assessment. These have been summarised in Table 4.4.

Table 4.4 Indicators of effective assessment

Please see print copy for Table 4.4

Adapted from James et al., 2002, p. 9)

There can be several beneficiaries of appropriately designed assessment:

- teaching/learning and assessment may be aligned – a benefit for all participants;
- the student, in terms of the learning afforded by engagement in *authentic* tasks and informative feedback on that learning;
- the teacher, in terms of feedback on the success or limitations in their address of defined student learning;
- parents, as they receive informed feedback;
- student accreditation, based upon summative assessment;
- funding bodies, as student results on assessment provide performance indicators. (McInerney and McInerney, 1994)

The first three of these beneficiaries form the subject of this study.

4.2.1 Awarding marks and grades: summative assessment

McInerney and McInerney (1994) identify the fundamental difference between formative and summative assessment:

> … (In formative assessment) the **process of learning** is of central importance, and rapid feedback is given to students to facilitate their mastery of the material to be learnt. … Summative measurement … refers to the process of determining whether or not students have achieved the ultimate objectives that have been set up in advance. … the emphasis is on the achievement of the student rather than the process of learning … (McInerney and McInerney, 1994, pp. 456-457)

The shift from terminal examinations as the sole means of assessment at universities has meant that formative tasks have frequently been included in assessment summatively. Whilst this has increased our perspective of student achievement across a broad range of tasks, it has simultaneously given rise to other problematic issues: student workload; dominance of the grading over the formative role; the negative impact of normative comparison on the formative role (Black, 1999, Black and Wiliam, 1998). With all work counting in the final grade, continuous assessment has resulted in increased workloads for students. It has also increased pressure of subject coverage for teachers and this has been accompanied by the increased burden of marking. (Piper et al., 1996)

This increased workload for students has tended to encourage the *surface learning* approaches evident in students who seek 'examination clues'. (Black, 2000) The increased classroom assessment load for teachers has resulted in a greater emphasis being placed on lower order knowledge and cognitive skills that are more readily recognised in marking. This in turn has fostered rote and repetitive learning in students. (Johnstone, 2005; Black, 2000; Piper et al., 1996)

There is general consensus in the literature that comparative grading offers little information about learning to either student or teacher. It promotes *surface learning* and may in fact it may adversely impact on student motivation. Highlighting personal *learning deficits* by comparison with the distribution of the *marks* of the cohort, offers little by way of encouragement or process to bridge the gaps, and may even deliver a counter productive blow to self esteem. Only those on the upper edges of the norm are affirmed and encouraged, and they, by ‘definition’, have no gap to bridge! For further discussion of this issue, see paragraph 4.2.3.

Biggs (1998) took exception to Black and Wiliam’s (1998) exclusion of summative assessment from their broad literature review of assessment. He maintained that the use of summative assessment as formative may be effective but its effectiveness is contingent upon its *alignment* with the instructional objectives and its marking against those objectives:

> A condition is that the assessment is deeply criterion-referenced, incorporating the intended curriculum, which should be clearly salient in the perceived assessment demands. When that happens, you get al.igned instruction, where teaching to the test is exactly what you want because it is teaching the intended curriculum. …when the *summative* assessment is defining the parameters for the *formative* assessment, it does not seem helpful to confine a review of the one with the attempted exclusion of the other. (Biggs, 1998, p. 107-108)

At least here, Biggs (as a proponent of Mastery learning) would find himself in accord with Black and Wiliam! They have identified that the planned feedback associated with mastery learning is strongly linked to the positive learning gains that have consequent positive impacts upon motivation. They have, however, cautioned that more research is needed to establish the most essential aspects of providing effective feedback. (Black and Wiliam, 1998) Neither would Biggs be disturbed by Black and Wiliam’s negative perspective of norm referencing approaches to assessment which he too recognised as

> …very common practices, but they make criterion-referencing of higher cognitive level performances all but impossible. (Biggs, 1999, p. 69)

On the surface Biggs’ claim, as an Australian academic, might not prove too provocative to many academics in Australian universities. However, although many Australian examiners reported using criterion referenced marking, only 5% were willing to ‘share’ their criteria (Piper et al., 1996)! Comments such as

> In my experience, <u>very</u> few departments have a systematic statement of detailed objectives, aims and outcomes for each unit. (Anonymous in Piper et al., 1996, p. 83)

have led Piper et al. (1996) to conclude that

> …it appears less likely that examiners were genuinely engaged in criterion referenced marking… (Piper et al., 1996, p. 83)

4.2.2 Informing students: formative assessment

> … research into student learning has shown the powerful effect of assessment in determining the real "curriculum in action" as opposed to the espoused curriculum. (Nightingale et al., 1996, p. 7)

Biggs (1999) and Ramsden (1992) also observed that assessment defines the learning for students. The desired learning behaviours to be demonstrated in the assessment should then impact on the choice of the type of assessment (Biggs, 1999; Ramsden, 1992), with particular care given to matching of the level of knowledge and skill. Hargreaves and Grenfell (2003) also recommended that

> The assessment design should include not only <u>what</u> the students need to do, but also <u>how</u> they are to do it. (Hargreaves and Grenfell, 2003, p. 2)

They advised forms of assessment that reflect professional practice in order to promote development of the generic skills espoused as university graduate attributes: critical thinking; professional, oral and written communication forms; group interaction; and peer review. (Hargreaves and Grenfell, 2003)

A second consequence of an assessment driven curriculum, is that it affords the optimal opportunity for aligning teaching and learning. If summative assessment were to evaluate student learning against the defined learning outcomes, then the teaching/learning framework should be designed to achieve those outcomes. Such an approach sets up an inevitable link between summative and formative assessment.

Yorke (2001) has perceived

> The assessment of student learning in higher education is under-theorised. (Yorke, 2001, p. 117)

In particular, he has also acknowledged,

> There is little to be found in the assessment literature as regards theory relating to formative assessment. (Yorke, 2001, p. 117)

Black and Wiliam (1998) and Sebatane (1998), in his review of their study, acknowledged that there has been no universally accepted definition of formative assessment, although much work has been done in understanding the complex relationships between teaching, learning and assessment. Black and Wiliam expressed their hope that their review of the 250 publications identified by them as noteworthy might help

> … to see whether the theoretical and practical issues associated with assessment for learning can be illuminated by a synthesis of the insights arising among the diverse studies … (Black and Wiliam, 1998, p. 7)

The definition of formative assessment formulated for the purpose of Black and Wiliam's (1998) literature review of formative assessment. was

> … all those activities undertaken by teachers, and/or by their students, which provide information to be used as feedback to modify the teaching and learning activities in which they are engaged. (Black and Wiliam, 1998, p. 8)

Nevertheless, they acknowledged difficulty in isolating the effects on learning of formative feedback from the associated teaching behaviours involved in the detailed planning required to align all facets of teaching/learning/assessment. (Black and Wiliam, 1998)

For assessment to be formative the learner must:

- perceive a gap between his/her current knowledge/skill, and
- take action to bridge that gap. (Biggs, 1998; Black and Wiliam, 1998)

Formative assessment shifts the focus from what is taught to what is learned. (Nightingale et al., 1996) This can be achieved through appropriate feedback on continuous assessment throughout a session, or through the use of summative assessment such as a portfolio, graded in recognition of the *quality* of the learning demonstrated. However for the latter to provide effective feedback, it should be designed to promote *meta-cognitive* structures and reflective practice in students by generating an awareness of the desired learning outcomes. (Black and Wiliam, 1998; Biggs, 1998) Development of the ability to self-assess is fundamental to *lifelong learnin*g (Sadler, 1998), but it should be noted that

> While challenging assignments and informative feedback can greatly strengthen learning, they do not necessarily help students develop their capacity for self evaluation. (Nightingale et al., 1996, p. 7)

To achieve this, students must recognise defined outcomes, critically reflect on their own performance against these, and take action to achieve them on the basis of effective feedback from formative assessment. (Sebatane, 1998)

Sebatane (1998) also cautioned the impact of high stakes assessment on the reform of formative assessment. Black (2000) too, expressed such a concern and recommended a re-focus from the 'exam' back to the 'learning'. He has questioned both the validity and reliability of high stakes assessment and hence viewed it as providing a poor focus for student learning.

But there are further issues that need to be resolved for formative assessment to be effectively used to promote student learning. It is contingent upon teachers having some understanding of the theory that underpins it. Yorke (2001) claimed that

> Assessors typically do not have any substantial grounding in the theory (limited as it is) and practice of assessment. (Yorke, 2001, p. 117)

This was substantiated by Piper et al. (1996) in their review of examination practices and procedures in Australian universities. Seventy percent of the responding examiners indicated that they had not

> … had any formal training in the setting and marking of examinations. (Piper et al., 1996, p. 24)

Whilst their evident expertise in their associated disciplines stands acknowledged, many professionals in the tertiary sector have been poorly trained (or not trained at all) in developmental psychology, communication and teaching:

> Lecturers in higher education are, in general, more *au fait* with the structure and progression of their subject discipline than they are with matters of student development. (Yorke, 2001, p. 118)

This obviously renders structured and meaningful feedback to the students difficult. (Yorke, 2001)

4.2.3 Norm versus criterion based assessment

There are two models of assessment:

1. The *measurement* model

> … is designed to assess personal characteristics of individuals, for the purpose of comparing them with each other or the general population norms. … (reduces) performances to numbers on a scale … (and) assumes that the characteristic being measured is stable, …(and) normally distributed. (Biggs, 1999, p. 69)

2. The *standards* model

> … is designed to assess changes in performance as a result of learning, … (Its purpose is) not to identify *students* in terms of some characteristic, but to identify *performances* that tell us what has been learned, and how well. (Biggs, 1999, p. 69)

Thus, evaluation of student learning may involve comparison of individual achievement against the achievement of a cohort of students (norm referenced assessment) or against a set of pre-established criteria (criterion referenced assessment). Norm referenced assessment does not recognise shifts in achievement across different sessions by different cohorts. However experienced examiners may believe themselves competent to adjudicate over such discrepancies based upon their experience (Piper et al., 1996). Norm referencing practices have supported the notion that it is the examiner's brief to 'spread' the students for grading. (Black, 2000; Biggs, 1999) Hence their aim has been to identify differences. Biggs (1999) argues persuasively that 'good teaching' should aim to reduce these differences not aim to increase them. In attributing *surface approaches* to learning on poor *alignment* of assessment with the objectives, he ascribed blame to

> … such institutional policies … requiring assessment results to be reported in percentages or "marks" or requiring results to be distributed along a predetermined curve. (Biggs, 1999, p. 69)

He further declared that there

> … can be no *educational* justification for grading on a curve. (Biggs, 1999, p. 69)

Criterion referenced assessment on the other hand, references students' performances against a defined standard - a more appropriate focus, particularly for summative achievement (Biggs, 1999). However criterion referencing frequently fails to eliminate the impact of an individual examiner's perceptions and values. (Piper et al., 1996) It is the opinion of the teacher/researcher, that clearly defined criteria in a criterion referenced assessment task, afford the opportunity to promote the *meta-cognitive* skills of self-assessment and monitoring against the expected outcomes, by providing the guiding mechanism.

Demands for national evaluations of student learning using highly standardised tests may indicate nothing more that that students can be well prepared to sit these tests and little more. These tests rarely cater for differences in culture and language. The may also fail to capture achievement because

> … human capacity can be displayed in different ways in different contexts and backgrounds, … (Black, 2000, p. 412)

Observed improvements in performance in standardised tests over initial years of implementation can be directly attributable to teachers accommodating time for 'teaching to the test'. Such approaches have arisen in the past in America and in the United Kingdom. (Black, 2000) In Australia, there have been increasing calls for 'national standards', but these have been associated rather with standardised curricula rather than the broad generalised tests such as the SAT in the United States.

4.2.4 Assessing specific learning: critical thinking

The recognition of the importance of the generic skills of critical thinking, problem solving, professional communication, information literacy and group interaction has resulted in universities defining them in their desired graduate outcomes. Some universities provide support 'skills' courses to students, but because of the complex interaction of discipline knowledge and the skills, the notion of divorcing the skill from the discipline concerned has been called into question (Black, 2000; Pellegrino et al., 1999)

> … the evidence now seems to suggest that human competence in any intellectual discipline has to be developed and measured in the context of that discipline. (Black, 2000, p. 413)

Hargreaves and Grenfell (2003) maintained that

> The use of assessment as a learning motivator is one way of teaching generic skills without jeopardizing the learning of discipline skills. (Hargreaves and Grenfell, 2003, p. 2)

Although these desired skills may be defined as graduate outcomes, they are not always included in subject outcomes, and even if included, are frequently not addressed directly in the instructional framework. Many academics believe them to be implicitly

> … buried deep within our teaching philosophies. (Hargreaves and Grenfell, 2003, p. 1)

In some subjects, these skills may have been assessed but student performance on tasks requiring them may have presented disappointment to academics. Such an experience at the University of Western Sydney has been addressed in Chapter Seven. In this study critical thinking was regarded by the teachers as including evaluation and integration of relevant information to inform judgement. (Te Wiata, 1996) Students in a final year subject were consistently failing to demonstrate such thinking. Academics were determined to develop a teaching/learning framework that would foster the skill development.

In other classrooms, the teaching/learning environment may focus upon the desired skill development, but assessment may target lower level skills. If students regard the assessment as defining the curriculum, they may circumvent involvement in the more demanding learning tasks that target the complex skills. This may be true even if the assessment is modelled on tasks requiring higher order skills! The skills should not only be engaged in assessment tasks, but also 'marked' if students are to appreciate the weight given them by academics and if students are to adopt a *deeper* approach to learning.

Tasks that

- model professional practice;
- are experiential;
- use collaboration and
- are problem-based

afford opportunities for demonstrating higher order thinking. However complex thinking may require modelling or other form of scaffolding as students develop the cognitive skills. The difficulty then arises that the assessment tasks

> … must be sufficiently dissimilar from those previously attempted as learning exercises to test real achievement rather than memory and regurgitation… (Sadler, 1998. p. 81)

Tasks that encourage *meta-learning*, such as portfolios, are believed to facilitate transfer of the skills of critical thinking and judgement beyond the classroom and promote a self-regulatory approach to *lifelong learning* (Hernandez, 2000; Biggs, 1999; Biggs, 1998; Sadler, 1998).

4.2.5 Assessing learning in statistics

delMas (2002) has argued that

> … an objective that is not assessed really is not an objective of the course … (and) the claim that this learning is a goal of instruction … seems to be a shallow one unless that learning is assessed. (delMas, 2002, p. 2)

delMas (2002) reviewed three articles by Rumsey, Garfield and Chance, noting the similarities and differences in their discussions of the goals of instruction in statistics. One pervading idea identified in these discussions was a caution

> … not to assume that understanding, reasoning or thinking will simply come in and of itself without making these objectives explicit to the student. … (and hence) … that we (must) not only make our objectives clear, but … follow through on these objectives by planning instruction to develop these outcomes and assessments that require students to demonstrate their understanding, reasoning, and thinking. (delMas, 2002, p.1-2)

Statistical learning was discussed in Chapter 3 (see paragraph 3.4). That paragraph addressed current trends in statistical education, and linked the domains of statistical literacy, reasoning and thinking. The educative processes required to inspire learning in these domains has been strongly linked to the practice of the expert. Essentially this necessitates a problem-based approach, supported by collaboration with peers and led by a practitioner. (Chance, 2002)

Researchers in statistical thinking (Chance, 2000; Moore and Cobb, 2000; Rossman and Chance, 1999; Cobb and Moore, 1997; Gal and Garfield, 1997) have displayed much in common with Hargreaves and Grenfell (2003). They too have sought to promote higher order thinking contextualised within the discipline knowledge of statistics. In designing assessment of this type of thinking, Chance has recommended

> …"assess what you value." If you are serious about requiring students to develop … (the habits of the statistics practitioner) … then you must incorporate follow-up questions into your assessment instruments, whether final exams or performance assessment components. (Chance, 2002, p. 7)

She personally has rewarded students

> … as much for the process as the final product … (allowing) … students to analyse data using the techniques discussed in the course … (and requiring

> discussion of) … potential biases and other critique of their own work, realize limitations of what they have learned, and see how theory differs from practice – all key components of statistical thinking. (Chance, 2002, p. 8)

Detection of achievement of these skills often has proved elusive in examinations. There is a need for more open ended questions and subsequent evaluation of students' approaches to solving problems and their ability to target the fundamental issues arising in them. (Chance, 2002)

delMas (2002) has argued that

> … defining the student behaviour that exemplifies a learning objective provides the impetus for instructional design. If assessments are then derived from the instructional experiences, students can form valid expectations of how their understanding will be assessed. Assessments tied to the objectives through instruction should be both meaningful and useful to students. (delMas, 2002, p.3)

He has reflected on how it might be possible to differentiate the goals of statistical literacy, reasoning and thinking, highlighting that it is not so much a question of content

> … but rather what we ask students to do with the content. … it is the nature of a test item that determines which of the three domains is assessed and possibly allows for more than one domain to be assessed by the same item. (delMas, 2002, p. 5)

In the past much of assessment in statistics focused on *doing* statistics, requiring student organisation and recall of the *content* of the course with little or no reference to demonstration of *contextualised understanding.* Most statisticians might have acquired expertise not as outcomes of academic pursuit but rather through years of professional practice. Elements of such practice incorporated into instructional strategies might prove more engaging for students and provide a more appropriate apprenticeship than subjects grounded in content and facility with calculations alone. (delMas, 2002)

4.3 Constructing learning

The aim in university teaching is to promote *deeper learning* in students in order to fit them for the demands of *life-long learning* expected by their professional and civic engagements. Biggs (1999) claims

> *Good teaching is getting most students to use the higher order cognitive level processes that the more academic students use spontaneously.* Good teaching narrows the gap. (Biggs, 1999, 58)

An active, problem-based teaching/learning environment is more likely to stimulate most students to engage such thinking (Biggs, 1999).The process of active participation in solving relevant problems, relating those solutions to theory and testing the knowledge in other contexts relates closely to learning resulting from life experiences. Impacting on this dynamic process are the physical and social environments, and histories and interests of the learner. Selection of appropriate teaching strategies requires a melding of all of these ideas. The instructional design should offer active learning, collaborative participation, defined and achievable goals and opportunities to reflect on the learning process itself.

4.3.1 Collaborative learning

> The ability to collaboratively create, write, and manage tasks and projects is becoming increasingly important in the business world. (Pfaff and Huddlestone, 2003, p.37)

The current belief in the value of team skills is reflected in specifications of desired graduate attributes at universities throughout the world. However, the value placed on such skills by business and communities is not the only driving force behind collaborative learning practices in universities. Livingstone and Lynch (2000) identify some of the key issues leading to their increased profile:

- Expository teaching/passive learning has been increasingly replaced by practices which actively engaging students in more socially oriented learning contexts;

- Expansion of learning outcomes beyond acquisition of discipline-based knowledge and skills to include personal development and social awareness;
- Group work is accepted as conducive to development of creativity and social skills; and
- Changing demographic patterns have resulted in increased student enrolments, no commensurate rise in funding for resources and resultant increases in class sizes, also confirmed in Australian Bureau of Statistics publications (Australian Bureau of Statistics, 2007) and affirmed by West (1998) and James et al., (2002).

A constructivist approach to tertiary education recognises the need for active engagement, social interaction and context for student learning, and collaborative learning structures can provide these (Hernandez, 2002; Watkins et al., 2002; DeMulder and Eby, 1999). Research has indicated the capacity of teamwork to improve individual performance (Pfaff and Huddlestone, 2003; Livingstone and Lynch, 2000; Johnson et al., 1998). Johnson et al. (1998) claimed that teamwork promotes persistence, transferability of knowledge, social skills and intrinsic motivation. Watkins et al. (2002) also included improved management skills, communication skills and multiethnic relations among its benefits. They believed that collaborative learning frees the teacher to allow them to be

> … more concerned with 'high level' enquiries and freed from mundane tasks. (Watkins et al., 2003. p. 5)

Hernandez (2002) also stressed its importance in the promotion of higher order thinking and *deeper learning.* Biggs (1999) also argued that students be able to

> … work collaboratively and in dialogue with others, both peers and teachers. Good dialogue elicits those activities that shape, elaborate, and deepen understanding. (Biggs, 1999, p.61)

Teamwork affords students opportunities to view problems from multiple perspectives (Hernandez, 2002, DeMulder and Eby, 1999) and can encourage a critical reflection that impacts upon students' value systems and senses of social responsibility (Livingstone and Lynch, 2000).

Johnson et al. (1998) perceived cooperative learning as relying on shared goals for achievement and hence not competitive and focused on the individual. They believed that

> Cooperative learning is the heart of problem-based learning. (Johnson et al., 1998, p. 28)

For groups to function effectively, Johnson et al. (1998) suggested the following features as essential:

> … positive interdependence, individual accountability, promotive interaction, social skills and group processing. (Johnson et al., 1998, p. 29)

Livingstone and Lynch (2000) explored how this might be achieved. They claimed that group dynamics were expedited by students 'applying' for the various identifiable jobs necessary to the set task such as project manager, project coordinator, graphic designer etc and group allocations being made on this basis. It also appeared to promote a perception of control over the group's effort not evident in groups which were randomly allocated.

Although collaborative learning might ostensibly decrease the demand on teaching resources, Kuh (2003) has cautioned that careful planning is required if students are not to be given opportunities for reduced effort in producing collaborative output. He described the potential for the 'disengagement compact' in which the student might be freed the interference of the teacher and vice versa. For effective learning, each aspect described by Johnson et al. (1998) should be addressed either through group preparatory classes and/or assessment devices which track both group and individual progress.

Livingstone and Lynch (2000) and Pfaff and Huddlestone (2003) explored the perceived negative side-effects of collaborative learning:

- unfairness in reward for effort for exceptional students;
- the 'free ride' aspect;
- 'hidden' laziness;

- greater demand on student time; and
- inequities in team allocations.

Although they detected a significant correlation between individual and group marks, it was not strong. It is apparent that other skills come into play in group work!

The studies of Livingstone and Lynch (2000) indicated that generally individual achievement (based on individual assessment marks) was lower than group achievement. Notable exceptions were the highest performing individual students. They also found evidence in support of the 'free ride' aspect, as the lowest performing individual students received top group marks. Student surveys nevertheless generally demonstrated positive experiences, but highlighted group 'protection' of poorly committed team members. Pfaff and Huddlestone (2003) also cautioned policing of the 'free ride'. They suggested *peer evaluation* for this purpose. Peer evaluations could then be used to moderate group results. The work of Livingstone and Lynch (2000) highlighted the need to monitor individual accountability, and for provision for assessment of team skills or at least to diminish their impact upon high achievers.

Team work provides opportunities for critical evaluation, judgement, mentoring and advising. (DeMulder and Ebe, 1999) Indeed if these are valued skills and espoused as desired learning outcomes in universities, then perhaps it is incumbent upon teachers to implement appropriate assessment strategies to target them.

4.3.2 Technology

Facility with technology is another espoused learning outcome in universities. With an increasing rate of growth in knowledge, technology affords speedy and widespread access to it. It also facilitates rapid communication. (West, 1998) These features can be used to advantage in education. Students in Australian universities are very familiar with technology, both in their academic and social lives. As a familiar environment, it can therefore provide an appropriate medium for assessment and feedback on that assessment, and a potential improvement in workload for heavily burdened academics. These advantages however may be accompanied by the undesired effect of increased opportunities for plagiarism. (James et al., 2002)

In this study, technology has not only been used to facilitate dialogue between students and the teacher and to expedite access to learning resources, but also to facilitate the processes of learning the discipline content. Students in statistics classes have actively engaged in simulation exercises, and computer calculation freed students from facile calculations, thus enabling them to focus on the more complex discipline thinking required in analysing real data.

4.3.3 Authentic tasks

The literature offers no clear cut definition of authentic tasks, and research evaluating their effectiveness is sparse. There are however, many reports of positive implementation in the classroom, and this researcher believes that the theory of experiential learning underpins their implementation. Such tasks involve practical application of theory and the cognitive demand can be augmented by inclusion of aspects addressing the *meta-learning* described by Watkins (2002) and illustrated in Chapter Three (see Figure 3.1). Authentic tasks may thus be used to challenge students and promote higher order thinking.

Authentic tasks require students to work on ‘real life’ problems which are both meaningful and relevant to them. They afford opportunities for students to construct meaning through engagement in direct application of discipline knowledge and skills in context. They may be experiential and provide for collaboration. Montgomery claimed

> Authentic tasks are often multidimensional and require higher levels of cognitive thinking such as problem solving and critical thinking. (Montgomery, 2002, p. 35)

In emulating a ‘real life’ challenge, these tasks address both the process and product of contextualised learning. (Montgomery, 2002) As with most professional disciplines, although the fundamental knowledge is essential to professional practice, so too is the adaptability required to transfer the knowledge and skills to practice. Experiential learning on authentic tasks models professional practice. James et al. (2002) claimed that

> Students respect assessment tasks they believe mirror the skills needed in the workplace. (James et al., 2002, p. 10)

If we aim to maximise the engagement of students in their learning, then use of authentic tasks may afford an appropriate step in that direction.

4.3.4 Active participation

In a constructivist learning environment, it expected that students will construct meaning through collaboration with teachers and peers while being actively engaged in solving problems. The importance of discipline knowledge and skills may be perceived to be paramount by many academics (Yorke, 2001; Piper et al., 1996) but student learning of these does not appear to be optimised by a transmission model of teaching that views students as passive recipients. This is especially true of less motivated students. Biggs claimed that actively engaging students on tasks that demand higher order cognition, such as problem-based learning tasks, bridges the achievement gap between well motivated and academic students and poorly motivated students focused on surface learning approaches. (Biggs, 1999)

4.3.5 Motivation

An instructional design is generally judged effective if it achieves its goals and is economical

> … in the use of instructional time, materials, and other resources. It is not generally viewed as relating to the motivational aspects of instruction except in a negative way. If an instructional event makes inefficient use of time and resources it can be boring or irritating to the audience. But, efficiency of delivery does not add to students' intrinsic interest in the situation. (Keller, 2007)

However, student achievement may not be driven by any motivational merit inherent in the instructional design, but because of extrinsic factors such as long term employment goals, fear of failure, prestige etc. Keller (2007) claimed that unless consideration is given to motivational factors in instructional design, only those already positively predisposed will be engaged. The challenge is to motivate those students not motivated to learn!

> It is true that academic achievement is greatest when pupils manifest felt needs to acquire knowledge as an end in itself. Such needs, however, are not endogenous but acquired – and largely through exposure to provocative, meaningful and developmentally appropriate instruction. (Ausubel, 1968, p. 11)

Biggs also maintained that

> “Motivation” is a product of good teaching, not its prerequisite. (Biggs, 1999, p. 61)

Keller’s ARCS model (2007) was developed around the key factors influencing motivation to learn:

- Attention: engagement of students in tasks that they believe to be valuable and interesting;
- Relevance: clearly defined outcomes that students perceive as meaningful;
- Confidence: clearly defined learning paths together with opportunities to succeed and effective feedback;
- Satisfaction: effective feedback and tasks that enable contextualisation of their learning.

Keller maintained application of the model to instructional design should encourage the intrinsic motivation essential to *deeper learning.*

4.3.6 Marking criteria

It can be so much easier to mark mathematics ‘by the pound’, a process that awards marks for steps in solving a problem, for each identifiable and relevant fact etc. But if the aim of assessment is to evaluate student achievement of the learning outcomes, designing appropriate assessment tasks may be only one step towards this end. The marking criteria also need to be matched against the objectives in such a way as to recognise all desired learning, including both discipline knowledge and the generic skills essential to its transfer and application in other contexts (James et al., 2002).

The marking criteria can also provide a measure of quality control for the marker, by clearly defining levels of competence and similarly, a transparent indicator to students of ‘fairness’ in recognising achievement/non achievement. (James et al., 2002) A sample set of objectives for an assignment assessing a linear relationship between two quantitative variables and an aligned set of marking criteria are presented in Tables 4.5 and 4.6 respectively. The strong relationship between the *scaffolding* illustrated in Table 4.3 should also be noted. The strategy of alignment of the learning objectives, the tasks and the marking of the tasks has been devised as part of the overall *constructive alignment* of the teaching/learning/assessment framework. (Biggs, 1999)

Table 4.5: Sample objectives for an assignment assessing a linear relationship between two quantitative variables

Objectives:

In this assignment students will

1. demonstrate an understanding of the difference between **data exploration** and **exploring relationships between quantitative variables**
2. produce regression analysis output from the specified statistical package
3. informally comment on a fitted linear model
4. explain the meaning of the correlation coefficient
5. determine the linear equation for the model
6. evaluate the significance of a linear relationship
   - referring to R^2
   - referring to the parameter estimates
   - referring to the ANOVA output
7. discuss the limitations of the model
   - with respect to the domain and range
   - the variation
8. form evidence-based conclusions about the appropriateness of the model

Adapted from Stat151/252 Laboratory Notes Week 10, 2005, Morris

In this example, the researcher has elected to deconstruct the desired knowledge and skills into their most basic components. A dichotomous marking scale has been assigned indicating achievement/non achievement. This reflects the researcher’s value for ‘knowing’ as opposed to ‘not knowing’ the fundamentals of the requisite discipline skills and knowledge. Her belief that imperfect fundamentals may present difficulties performing more complex cognitive tasks experience is predicated on years of experience.

Table 4.6: Example of marking criteria for an assignment assessing a linear relationships between quantitative variables

Criteria	Mark	Achieved/not achieved
1. ***Explore*** the potential for a relationship		
Construct a scatter plot	1	
Check for non-linearity	1	
Comment on **strength and direction of the relationship**		
From the scatter plot	1	
Determine the value of the correlation coefficient	1	
Explain the significance of the value of the correlation coefficient	1	
In terms of its size	1	
In terms of its sign	1	
Check for outliers/influential points	1	
Check for alternative models	1	
Check for patterns	1	
Total Mark 'Explore'	10	
2. ***Determine*** the form of the relationship		
Identify the dependent and independent variables	1	
Find the **equation** of the line of best fit		
Determine the value of the intercept	1	
Determine the value of the slope	1	
Write the equation for the relationship between the two variables	1	
Total for 'Determine'	4	
3. ***Evaluate*** the significance of the relationship		
Determine the value of the coefficient of determination	1	
Explain the significance of the value of the coefficient of determination	1	
Examining the t tests on the parameter estimates		
constant	1	
slope	1	
Analysing variation (ANOVA)		
Explained by the model	1	
Not explained by the model	1	
The F statistic	1	
The p value for the F test	1	
Total for 'Evaluate'	8	

Table 4.6 (continued): Example of marking criteria for an assignment linear relationships between quantitative variables

Criteria	Criteria	Criteria
4. ***Discuss*** the **limitations** of the relationship		
The domain of values	1	
The variation	1	
The model structure	1	
Total for 'Discuss'	3	
5. ***Conclude*** about the model based on the examined evidence		
In the **context** of the problem and including **evidence** from:		
Results of '**exploration**'	1	
Results of '**evaluation**'	1	
While recognising the **limitations**	1	
Stating the '**determined**' form of the relationship	1	
Total for 'Conclude'	4	
Total for task	29	

Implementing such an approach affords the potential for students to reproduce by rote its 'pre-processed' thought patterns. To mitigate this, tasks should stimulate sufficient challenge for transfer of the requisite knowledge and skills to distinct contexts.

> … tasks must be sufficiently dissimilar from those previously attempted as learning exercises to test real achievement rather than memory and regurgitation (unless that is the legitimate aim). But they must also be similar enough to fall within the region that reasonably allows transfer or extended application of learning. (Sadler, 1998, p. 81)

Section 5 of the marking criteria in Table 4.6 has been included in recognition of a student's ability to critically evaluate the evidence accumulated throughout the first four sections to form a contextualised conclusion.

The literature reveals a widespread application of *rubrics* in the marking of assessment. Common to most studies is a mark/grade allocation based upon comparisons between bands of performance descriptors indicative of difference levels of achievement. Application of such rubrics, however, cannot eliminate a subjective element in the marking process. The NSW Board of Studies applies such a marking regimen in essay/ written response type of examinations, but supports it with rigorous checks involving 'check marking' of sampled student papers and examination of marker statistics to maintain comparability (Masters, 2002). At university there has been little evidence of this approach in mathematics or statistics examination marking (Piper et al., 1996) although project work in these disciplines may have been commonly assessed in this way.

4.3.7 Feedback

Closely aligned with the notion of formative assessment, feedback was introduced in Paragraph 4.2.2. The discussion here relates particularly to the approaches to feedback used in this study.

Black and Wiliam (1998) identified four key elements that constitute systematic feedback:

- data on the actual level of some measurable attribute;
- data on the reference level of that attribute;
- a mechanism for comparing the two levels, and generating information about the gap between the two levels;
- a mechanism by which the information can be used to alter the gap. (Black and Wiliam, 1998, p.48)

The notion of 'measurable' might appear surprising as it connotes a quantification that is more closely aligned with the grading and/or norm referencing identified as a weakness in current approaches to assessment practice; (Black, 2000; Biggs, 1999; Black and Wiliam, 1998; Biggs, 1998) However the primary focus is not the 'measurement' per se, but the identification of the gap between 'actual' achievement and some 'reference' level, whether it be a prescribed mark, normative scale, descriptor or band of descriptors, or a criterion.

Effective feedback then must be timely, task related and standards oriented, (Sadler, 1998; Crooks, 1988) and interpretable for the student (Black, 1998). If the ultimate goal of teaching is to foster *lifelong learning*, then the aim must also be to relinquish the role of assessor to the student themselves. Appropriately framed feedback can also promote self regulation as

> … some of what the teacher brings to the assessment act must itself become part of the curriculum for the student, not an accidental or inconsequential adjunct to it. (Sadler, 1998, p. 82)

Sadler concluded that only a

> … framework … separated from cohort performance … (can)… remain stable so that real (or absolute) improvement can be plotted for each student. (Sadler, 1998, p. 84)

and that quality feedback should incorporate an understanding of

> … not just the technical structure of the feedback (such as its accuracy, comprehensiveness and appropriateness) but also its accessibility to the learner (as a communication), its catalytic and coaching value, and its ability to inspire confidence and hope. (Sadler, 1998, p. 84)

In this study, Ausubel's concept of an *organiser* has been used to frame models of statistical thinking to facilitate organisation of connected concepts and procedures. Assessment tasks were aligned with all other facets of teaching and learning through the specified learning objectives. Detailed descriptive *scaffolding* underpinned the development of marking guides that were designed to inform and organise student responses. These guides were used as the only criteria for evaluation. They were 'checked' and returned to students as detailed feedback. The marking guides thus

- defined the expected knowledge or skill to be marked;
- indicated achievement/non achievement of the required knowledge or skills;
- identified the 'gap' between 'reference level' and 'actual' achievement and
- afforded a mechanism for improving their performance.

4.3.8 Learning portfolios

Definitions of learning portfolios encompass the notion of the student's learning journey illustrated through collections of the student's work and incorporating reflection and critical analysis of

> … not only their own academic and skill development but also to reflect on the teaching and learning experience. (McElwee et al., 1992, p. 29)

It is the student who accumulates and comments upon its content and 'ownership' of the learning is paramount (Klenowski et al., 2006; Liu et al., 2004; McElwee et al., 2002) if it is to effect the maximum skill development. The primary benefit of the learning portfolio is the developing awareness in the student of the process of their own learning. Thus the portfolio provides evidence of *meta-cognition*, and it is such skills that will foster the self-regulation that enables support of future learning and promotes appreciation of the skills and knowledge of others. (Klenowski, 2006; Biggs, 1998)

This form of assessment requires clear definition of desired learning to facilitate assessment of the portfolio, and teacher support if students are to approach the task with confidence and achieve the high order cognitive outcomes it is designed to address.

The learning portfolio has been used in the case studies discussed in this study (see Chapters 6 and 7). However the notion of portfolio has been somewhat 'stretched' in the statistical learning study (see Chapters 6) to include the student 'workbook'. Although no reflective component was required in its completion, students were encouraged to develop 'frameworks' to organise and summarise the learning. These frameworks were intended to permit generalisations that would facilitate transfer and application of the discipline knowledge and processes. These applications are further discussed in the specified chapters.

4.3.9 Peer evaluation

Critical evaluation and self-reflection are essential skills for supporting *lifelong learning*. They involve the highest order cognitive skills of judgement and *meta-cognition*. Peer evaluation is an approach to assessment that addresses these skills. Its success however is also consequent upon an assessment framework that

- has clearly defined outcomes;
- promotes active and collaborative learning;
- enlists the continuous support of the teacher;
- values effective feedback;
- uses authentic tasks;
- provides explicit marking criteria and trains students in application of those criteria. (Juwah, 2003)

Involvement of students in setting the criteria strengthens student motivation and commitment. (Juwah, 2003; Toohey, 1996) Toohey (1996) has also recommended moderating self-assessment using peer review.

In the case study discussed in Chapter 7, peer review was used in tutorial exercises to acquaint students with the marking criteria used to mark the essay and presentation assessment tasks. Through exposure to and considered evaluation of other students' work against the criteria, students were given opportunities to demonstrate a more personal dimension of the key skills of the evaluative and critical thinking targeted by the subject.

Chapter 5
Evaluation
Learning from teaching: Looking for atoms of meaning

The White Rabbit put on his spectacles. “Where shall I begin, please your Majesty?” he asked.

“Begin at the beginning,” the king said gravely, “and go on till you come to the end; and then stop.”

There was dead silence in the court, whilst the White Rabbit read out these verses:-

“They told me you had been to her,
And mentioned me to him:
She gave me a good character,
But said I could not swim.

He sent them word I had not gone,
(We know it to be true):
If she should push the matter on,
What would become of you?

I gave her one, they gave him two,
You gave us three or more;
They all returned from him to you,
Though they were mine before.

If I or she should chance to be
Involved in this affair,
He trusts you to set them free,
Exactly as we were.

My notion was that you had been
(Before she had this fit) An obstacle that came between
Him, and ourselves, and it.

Don’t let him know she liked them best,
For this must ever be
A secret kept from all the rest,
Between yourself and me.”

“That’s the most important piece of evidence we’ve heard yet,” said the king rubbing his hands; “so now let the jury – ”

“If any one of the can explain it,” said Alice, (she had grown so large in the last few minutes that she wasn’t a bit afraid of interrupting him,) “I’ll give him sixpence. I don’t believe there’s an atom of meaning in it.”

The jury all wrote down on their slates, “*She* doesn’t believe there’s an atom of meaning in it,” but none of them attempted to explain the paper.

“If there’s no meaning in it,” said the King, “that saves a world of trouble, you know, as we needn’t try to find any. And yet I don’t know,” he went on, spreading out the verses on his knee, and looking at them with one eye; “I seem to see some meaning in them, after all. ‘- *said I could not swim* - ,’ you can’t swim can you?” he added turning to the Knave.

The Knave shook his head sadly. “Do I look like it?” He said. (Which he certainly did *not*, being made entirely of cardboard.)

From “Alice’s Evidence” in
Through the Looking Glass by Lewis Carroll, pp. 104-105.

5.0 Reality in the qualitative study

> Qualitative researchers stress the socially constructed nature of reality, the intimate relationship between the researcher and what is studied, and the situational constraints that shape inquiry. Such researchers emphasise the value-laden nature of inquiry. They seek answers to questions that stress *how* social experience is created and given meaning. (Denzin and Lincoln, 2003, p. 13)

There are many perspectives of reality within the classroom. Each is tinged by the needs, motivations and histories of the observers: the teacher; the students; and the administrators. To evaluate the quality of student learning requires aggregation of these perspectives to refine any picture of the outcomes of teaching innovations. It is important to note, however, that not only has the researcher participated in the action, but it is her hand that has guided the formulation of the evaluation framework for judging the effectiveness of the innovations. Indicators of their effectiveness in promoting student learning have arisen from both quantitative and qualitative sources. The researcher has endeavoured to include individual perspectives in her analysis, provide rich descriptions of the classroom contexts and acknowledge the constraints arising from the conflicting needs of the participants (Denzin and Lincoln, 2003). Thus multiple data sources and perspectives have melded to give a single, more comprehensive view of the action.

The quantitative aspects of this study have not therefore reflected any adherence on the part of the researcher to the positivist or post positivist traditions. The researcher has not been driven by an ultimate reality, knowable or otherwise. For this study, the defining features of the classroom reality have reflected her own contextualised perceptions of effective practice, with the criteria for judgement influenced by her *value-laden* understanding of what constitutes 'desirable' student learning (Gomm, 2004). Rather, the quantitative data have provided indicators in her evaluative framework.

Because this study has been framed in the knowledge, experience and values of the researcher, it has been important, and certainly useful, to clarify some aspects of her philosophical position before construction of an evaluation framework. As a pragmatic researcher, her ontology has relied upon the notion of effectiveness (Mertens, 2005). Her criteria for judgement have arisen from a personal perspective, albeit reinforced by peers and grounded in the literature. She has declared her epistemology to be *value-laden*. Her methodology has utilised all available sources of data (Mertens, 2005, O'Leary, 2005).

5.0.1 Situating the researcher

In undertaking this study, the researcher's primary intention was to implement a learning framework that enhanced student learning and to systematically evaluate the practice. The focus on systematising the evaluation was an addition to her usual reflective and responsive approach to her practice and it has heightened her alertness to confrontations of the protagonists in the classroom and the changes and stalemates that resulted from them.

Thus informed by a body of educational theory and research, reinforced by years of educational practice, and armed with a modicum of 'expert' discipline knowledge, the researcher has approached this study from a pragmatic ontological position. For the researcher, the 'reality' of her classroom has been 'what works' for student learning, that is 'effectiveness' has provided the gauge of successful teaching (Mertens, 2005) Defining what 'works', however may differ depending upon the perspective of the teacher, and it is the teacher who ultimately sets the criteria for what 'works'. Hammersley (2004) cautioned that, far from objective judgement of effectiveness,

> … such assessment cannot be separated from value judgements about desirable ends and appropriate means; not without missing a great deal that is important. (Hammersley, 2004, p. 137)

The researcher openly acknowledges her imprint upon the definition of the classroom reality. Her 'reality', however, has been cross-referenced with those of other participating teachers. Her notion of 'effectiveness' has been founded in theory and the literature, forged through experience and moulded by continuous exchange with peers.

Throughout her teaching career, the researcher has remained strongly committed to a *deep learning* encompassing the *meta-learning* that facilitates self-regulation and promotes lifelong commitment to learning. This commitment has been reflected in the teaching/learning frameworks she has helped to devise for each of the case studies.

Because of their perceived value, these concepts have been further developed throughout this work.

The researcher's yardsticks for student learning have not been restricted to 'passing' assessment, but have included a measure of the level of student confidence in their own learning. She has always believed that 'good' students *know what they know* and *know what they don't know.* In her experience, such confidence has always resulted from a *deeper learning*.

Experienced teachers' observations have also provided indicators of general shifts in class learning across the years. The subject coordinators associated with this study have provided valuable insight through discussion and annotations in the researcher's journal.

In this study, the learning objectives for each subject (see the case studies in Chapters 6 and 7) were virtually prescribed by discipline and course requirements and hence moulded the desired outcomes for each context. This work contains no discussion of the appropriateness of these objectives except in justification of the researcher's clarification of the terms *statistical thinking* and *critical thinking*.

The research has encompassed two distinct case studies:

1. Exploring Variation and Uncertainty in Data (known as STAT131): a classroom context addressing the fundamentals of statistical thinking. This case study has tracked student learning in response to the designed pedagogical framework;
2. Introduction to Accounting Philosophies and Theories (200102): a 'final year' undergraduate subject aimed at developing the critical and evaluative skills of students in an accounting context.

In each case, the pedagogy was devised to address the knowledge and skills specified at discipline level and the more globally defined university graduate attributes.

5.0.2 What happened and how do we *know* it?

In order to incorporate multiple perspectives of the action in the classroom, the researcher has had to canvas her potential sources of 'knowledge':

- what she herself 'knows' from experience, theory and the literature;
- what she 'sees', both in terms of classroom action and assessment feedback;

- what others ‘see’;
- what others ‘say’ and
- what is done, by both the researcher and other stakeholders.

These data have required thoughtful examination and re-examination to facilitate interpretation. Through experience, wide literature review and peer consultation, the researcher’s understanding of student learning has conceptually developed throughout the study and hence so too has her perception of what has ‘worked’. This has also impacted upon the evaluation framework.

> The processes that define the practices of interpretation and representation are always ongoing, emergent, unpredictable and unfinished. (Denzin and Lincoln, 2003, p. 420)

Evidence that was found lacking in early implementations has later been collected as the researcher came to recognise further dimensions to her image of student learning. Refinements were also made to the teaching/learning frameworks as its deficits were revealed.

5.0.3 Trees from the wood? Chasing *atoms* of meaning

In professional practice, a teacher responds to student needs both as they arise and in consequence of reflection in an informal evaluation of the classroom action. Formalising this evaluative process has required identifying potential sources of evidence for each of the multiple perspectives. Those identified have been summarised in Table 5.1. The sources highlight the researcher’s mixed method methodological approach, using both quantitative (e.g. assessment) and qualitative data (annotated journal and student comments). Interpretation of the data requires referencing individual interests and cross-referencing the multiple perspectives.

Table 5.1: Multiple perspectives and potential evidence sources

Perspective	Potential evidence source
Researcher	Annotated journal Reflective practice
Students	Surveys Observations of behaviours Interviews Assessment Completion, submission and attendance rates
Assessment	Marks: one side of the assessment coin • assessment marks • grade distributions Questions, answers and criteria: the other side of the assessment coin • applying Bloom to questions, answers and criteria
Teachers	Learning framework construction/reconstruction Annotated journal Peer review
Aligned perspectives	Attendance rates/submission rates and overall student perceptions (‘fairness’, defined learning) Perceived achievement and marks

5.1 Program evaluation

Robson (2002) describes an evaluation in the context of social enquiry as

> *… an attempt to assess the worth or value of some innovation, intervention, service or approach.* (Robson, 2002, p. 202)

Many authors (Mertens, 2005; Denzin and Lincoln, 2003; Robson, 2002) have, however, highlighted the wide range of accepted definitions of ‘evaluation’. These same authors have also distinguished between evaluation studies and social science research. Evaluation studies frequently seek evidence of the effectiveness of social or educational programs, and because of this, the ‘voices’ of interested (and possibly politically motivated) parties may exercise a greater impact than a more objective approach to social science research. (Mertens, 2005) Thus this has given rise to

> … pluralist conceptions of evaluation in which multiple methods, measures, criteria, perspectives, audiences, and interests were recognized. (House, 1993, in Mertens, 2005, p. 46)

Robson claimed that

> Evaluations also highlight issues to do with change. (Robson, 2002, p. 203)

These aspects are particularly relevant in evaluation studies that promote empowerment of the participants through collaborative assessment of progress, or those that focus on the functional changes of organisations under program implementation. Despite the potential political aspects of evaluation studies, social science research and evaluation studies both provide for

- monitoring of the program's progress;
- setting of standards or criteria for evaluation;
- judgement of the program's merit against these standards/criteria or based upon achievement of the program's objectives. (Mertens, 2005)

Effective evaluations use systematically collected information to help defined users, make decisions about effectiveness, rectify inappropriate action and plan (Robson, 2002) The Joint Committee on Standards at the Evaluation Center at Western Michigan University has constructed a draft set of Program Evaluation Standards (2007). Descriptive statements have been provided for the key aspects:

- Feasibility: practical procedures; political viability; resource use and project management.
- Propriety: responsive and inclusive orientation; negotiation of needs and expectations; human rights and respect; balance of needs; transparency and open disclosure and mediation of conflicts of interest and fiscal responsibility.
- Accuracy: trustworthy conclusions and decisions; explicit reasoning; valid constructs; reliable data; sound statistical methods; sound designs

and analysis; trustworthy evaluand and context descriptions and valid reporting.

- Utility: evaluator credibility; attention to stakeholders; clear purposes; explicit values; relevant data; meaningful processes and products; responsive reporting and concern for consequences and influence. (The Joint Committee on Standards at the Evaluation Center, 2007)

In this study, the researcher has used evaluation theory to track effectiveness of subject teaching practice for student learning. The primary foci have been the students and the teachers, with minimal consideration accorded to other potential stakeholders. The theory has facilitated systematisation of both evidence gathering and analysis, but it is important to note that the evaluator in this context is also a teacher and hence a stakeholder. (Mertens, 2005)

Bain (1999) highlighted

> … the need for learning centred evaluations in higher education. (Bain, 1999, p. 165)

and recommended that scholarly articles founded upon innovations in teaching and learning should:

> - Indicate the ways in which the teaching/learning intervention is noteworthy and innovative;
> - Draw upon the student learning and evaluation literatures to frame and interpret the innovation and its impact;
> - Carefully examine the learning processes and outcomes targeted by the innovation; and
> - Provide sound evidence for the influence of the innovation on learning processes and outcomes. (Bain, 1999, pp. 165-166)

These recommendations helped channel the search for an appropriate evaluation model for this study.

5.1.1 Evaluation models

In selecting an appropriate evaluation model, the researcher was acutely aware of her constructivist learning contexts and the logistics of implementing pedagogy in large teaching environments. Four models for evaluating teaching practice were considered in her examination of potential models for this study, and they have been discussed briefly here to highlight the deliberations undertaken in selecting a framework for evaluation.

5.1.2 The Context Input Process Product Model

The Context Input Process Product model (known as the CIPP model) resulted from the collaborative efforts of Stufflebeam, Foley, Gephart, Guba, Hammond, Merriman and Provus. Isaac and Michael, 1990) The model is predicated on the following three points

> First, evaluation is a continuous, systematic process. Second, this process includes three pivotal steps: (1) stating questions requiring answers and specifying information to be obtained, (2) acquiring relevant data, and (3) providing the resulting information as it becomes available to potential decision makers ... Third, evaluation supports the process of decision making by allowing the selection of an alternative and by following up on the consequences of a decision. (Isaac and Michael, 1990, p. 6)

As the primary purpose of the model is support of decision making, it has been designed to be equally effective for planning (goal and objective specification), structuring (strategic design for objective achievement), implementing (engagement of the strategies and ongoing improvements) and recycling (program continuation, modification, or disengagement) decisions. To this end, the model addresses four types of evaluation:

1. ***Context*** evaluation that identifies the shortfall between desired outcomes and the *status quo*;
2. ***Input*** evaluation that identifies relevant strategies and resources;
3. ***Process*** evaluation monitoring the implementation of the selected strategies;

4. ***Product*** evaluation facilitating the overall effectiveness of the program in terms of its meeting the objectives. (Isaac and Michael, 1990)

This model can be simply adapted to the educational environments of this study. In each case study, the researcher undertook

- a needs analysis to identify learning deficits/needs,
- a learning framework design based upon experience and the literature to facilitate student achievement of desired learning outcomes,
- implementation (involving continuous and formative monitoring, and amendment of the selected strategies), and
- a summative evaluation of program achievement at the study's conclusion.

Much of the decision making undertaken was by the teaching teams involved, and the evaluator was independent of these teams only in the smaller of the two case studies undertaken at the University of Western Sydney case study (see Chapter Seven).

5.1.3 Metfessel-Michael Paradigm

This evaluation model (see Appendix 5.1) has targeted the impact of implementation of an educational program on the whole school. It follows linearly through eight distinct steps:

1. Involvement of the whole school community.
2. Construction of an hierarchical list of specific outcomes appropriately framed so that criteria for judgement can be formulated;
3. Design of a learning framework that is conducive to achievement of the desired outcomes.
4. Development of measures for judging the effectiveness of the program.
5. Continuous monitoring for changes relevant to the specified objectives.
6. Use of statistical analysis on observed data relevant to the specified objectives.
7. Interpretation of the evidence in terms of the criteria set for judgement of program effectiveness.

8. Making recommendations based upon the evidence for continuation, modification or cancellation of the implemented program. (Isaac and Michael, 1999)

Although much of this model is also easily adapted for this study, the researcher recognised she wanted the facility to recycle through the implementation and evaluation phases. This opportunity is lost in such a linear model. Neither does the model accord the necessary emphasis on the learning framework design. This step is complex and time consuming, requiring reference to the literature and evaluation of professional experience.

Isaac and Michael (1990) cautioned users of the model that there exists the potential for correlated gains/losses effected by influences other than the program (such as different teachers or factors arising from the general school community). As a teacher working in a constructivist environment, such occurrences have been expected, recognised and declared by the researcher.

5.1.4 A Four stage evaluation of teaching and learning

This model (see Appendix 5.2) was adapted from one developed by Robert Glaser (1962), extended by De Cecco (1968) and Metfessel and Michael (1967) and formalised by Michael (1969) (Isaac and Michael, 1990). It incorporates

A. FORMULATION OF A COMPREHENSIVE AND CONSISTENT SET OF BROAD GOALS AND SPECIFIC INSTRUCTIONAL OBJECTIVES
B. (Description of) ENTERING BEHAVIOURS (INTERNAL CONDITIONS) IN STUDENTS AND TEACHERS
C. (Description of) INSTRUCTIONAL PROCEDURES (TEACHER INFORMATION PROCESSING)
D. EVALUATION OF PERFORMANCE WITH FEEDBACK LOOPS TO STEPS C, B AND A. (key points of the model from Isaac and Michael, 1990, pp. 18-19)

Whilst this model would appear to meet most of the needs of this study, and broad descriptions of the cohorts of students have been provided for each of the studies, no detail has been given of entry behaviours for either students or teachers. The goals and objectives for the targeted subjects were specified for *all* students taught by *all* teachers and were formulated without reference to entry behaviours. In addition, neither remediation nor teacher instruction has formed a part of any of the normal subject regimen. Some associated discussion has been included but only where obvious discrepancies in observed performance were perceived as potentially in consequence of teacher/student entry behaviours. The others three stages (A, C and D) were nevertheless very relevant.

5.1.5 An evaluation model for educational innovations

Alexander and Hedberg (1994) sought to develop an evaluation model for technology-based innovations. Their model (see Appendix 5.3) addressed evaluation questions and methods at the design, development, teaching and institutionalisation phases. Because of the relationship to evaluation of learning interventions in the classroom context, this model offered practical and direct connection to this study. Bain (1999) adapted this model to facilitate review of papers evaluating teaching/learning innovations in higher education. His modifications extended applicability beyond those technologically-based. His claimed advantages of the model were

> - it presumes that evaluation will occur in each of the major phases of an educational project (design, development, implementation, and institutionalisation);
> - it outlines the types of evidence and methods that may be appropriate for each phase; and
> - it demonstrates how close attention to the learning process and outcome should be threaded through all phases of evaluation. (Bain, 1999, p. 166)

This project does not examine effects at the institutional level. It frames its evaluation of student learning within the context of the selected subjects with reference to student achievement of their specified learning outcomes. The remaining phase descriptions, however, exhibited direct relevance to the case studies in this thesis and hence this model appeared to afford the strongest potential for adaptation to this study.

5.2 Developing a model fit to purpose

Evaluation of this study, has demanded a framework incorporating the facility for evaluation of an educational program that has

- been founded on specified learning objectives;
- occurred in a context laden with the values of the researcher, teachers and associated discipline staff, and the students;
- given rise to evidence arising from both qualitative and quantitative sources; and
- required repetitive application of the *Development, Implementation* and *Review Phases* of the evaluation for the action research cycles in the STAT131 case study.

The search for the appropriate framework has required careful examination of the evaluation process itself in order to contextualise the researcher's specific needs. Mertens (2005) has suggested the following three steps in the evaluation process:

1. Focusing the evaluation;
2. Planning the evaluation and
3. Implementing the evaluation.

The first two steps have been covered in the following discussion, and the final step has been included in each of the case studies and again in the concluding chapter, Chapter Eight.

5.2.1 Logic behind the program: mapping perspectives

In order to evaluate the study, the researcher embarked on the complex task of describing the *evaluand*. This process proved to be the most arduous of the project. Even for an experienced and reflective teacher, describing what was to be evaluated forced holistic consideration of aspects of teaching that had been previously examined in comparative isolation. The researcher's perspective of student learning has developed throughout this study, not so much in terms of modification of an existing vision, but in terms of an increased awareness of complexity and dimension. This emergent understanding is evident in changes in evaluation across the term of the study, particularly in assessment design and student surveys. In order to describe the evaluand, the research has again drawn support from the literature.

Mertens (2005) described a logic model as a useful tool for describing the evaluand.

> … (it) is a graphic technique that allows the explicit depiction of the theory of change that underlies a program. It should reflect a model of how the program will work under certain conditions to solve an identified problem. Logic models can be quite, including three major components: the program that is delivered, the people who receive the program, and the results that are expected. Arrows connect the various elements … and indicate the logical flow from one element to the next. (Mertens, 2005, p. 65)

The researcher has chosen to adapt the notion of the logic model to facilitate the description of the logic flow throughout the teaching/learning frameworks underpinning the two case studies that form this investigation. The results of her deliberations have been generalised and illustrated in summary in Figure 5.1.

The 'Inputs' column of the diagram identifies the protagonists in the constructivist learning environments who impact upon student learning, together with the library of supportive resources arising from the programs residing on the Website. The 'Supported Learning' column identifies the features of the framework upon which these *inputs* have the greatest impact. The researcher has not been identified as a separate *input*, despite her involvement in construction of the teaching/learning framework. She has been illustrated rather as an agency for facilitation of the *flow* in the process, as her influence pervades all of its aspects.

In an acknowledged constructivist learning environment, it may seem negligent not to include students as *inputs*. However this study aims to augment the learning of *all* participating students and in its design attempted to accommodate different learning styles and academic and cultural backgrounds. As with most *real* classrooms, the learning environment has been adaptive, responding to individual and collective deficits and misconceptions as they revealed themselves.

Please see print copy for Figure 5.1

Figure 5.1: A logic model for the teaching/learning framework

Adapted from Mertens, 2005, p. 65

The programs under evaluation have encompassed the middle two columns of the diagram. Alignment of their various facets has been highlighted by the inclusion of *Objectives* in almost all of the groupings (with the *Website* facets representing collections from each of the others). Assessment was used as a driver for teaching and learning, and achievement in that assessment evidenced one dimension of student learning, the others provided by students' personal perceptions, staff perceptions and examination of the knowledge and skills demanded by the tasks themselves. The stakeholders, and hence the multiple perspectives of the classroom reality, are listed in the final column, with each grouping enumerating the expected outcomes of the program.

5.2.2 Focusing the evaluation

This study has sought evidence that specifically devised learning frameworks have impacted upon student learning in two distinct case studies. In describing what is to be evaluated, the researcher has been conscious of reported deficits in student learning from the teachers of all three subjects under scrutiny. Student achievement had consistently fallen short of teacher expectations, particularly in the University of Western Sydney case study (see Chapter 7). Statistical thinking is rarely taught before students reach higher education. Hence most students approach it with some trepidation. The statistical subject examined (see Chapter 6) was a compulsory unit, but many students failed to see its relevance.

Broad descriptions of the evaluands, the reasons for evaluation, potential stakeholders and possible constraining influences (Mertens, 2005) have been detailed in Table 5.2. Commonality is evident across the case studies and it is this that allowed the construction of common evaluation questions for all of them (see Table 5.3).

Table 5.2: Focusing the evaluation[1]

Focus questions	STAT131	Introduction to Accounting Theories and Philosophies
What is to be evaluated?	The impact of all facets of the teaching/learning framework generally on student learning as defined by the objectives and more specifically on statistical thinking.	The impact of all facets of the teaching/learning framework generally upon student learning as defined by the objectives and more specifically on critical and evaluative thinking.
Why evaluate it?	To ascertain the degree to which the aim of the program has been achieved, i.e. the enhancement of student learning by clear and transparent specification of the learning objectives supported by a specifically designed learning framework.	To ascertain the degree to which the aim of the program has been achieved, i.e. the enhancement of student learning by clear and transparent specification of the learning objectives supported by a specifically designed learning framework.
Are there stakeholders?	Researcher, teachers, students and faculty.	Researcher, teachers, students and faculty.
Are there constraints?	Researcher and faculty.	Researcher and faculty.

[1] The focus questions have been included in Table 5.3.

Source: Adapted from Mertens, 2005, p. 62-70

Table 5.3: Focus questions for developing the evaluation framework[1]

Description questions	Implementation questions	Impact questions
Does each facet of the teaching learning framework address the desired learning outcomes? Why is each facet to be included in the teaching/learning framework? Is its implementation similar for all students? Is teacher training/support required?	Was each facet of the teaching learning framework implemented as designed?	Were there changes made to any facet of the teaching learning framework across sessions during the study? What motivated the changes across the sessions during the study? How were decisions for change made? Were there differences in the desired outcomes: 1. in comparison to those observed prior to the study? 2. across sessions during the study? Were all students affected equally?

[1] These questions applied to the following facets of the teaching/learning framework in each of the three case studies:

- subject objectives;
- subject presentation;
- subject assessment; and
- subject evaluation.

Adapted from Mertens, 2005, p. 62-70

5.2.3 Evaluation criteria: the search for *atoms* of meaning

An issue facing evaluation of classroom innovation is the plethora of impacting variables. Alexander and Hedberg (1994) identified Tucker's (1993) use of a *checklist* in constructing an evaluation instrument as a potential solution to this dilemma. (Alexander and Hedberg, 1994) Although *checklists* do not eliminate the possibility of the *subjective* influence in selection of items for inclusion on the list, checklists may afford a more *objective* opportunity for judging merit.

Scriven (2005) and Stufflebeam (2000) also promoted the use of checklists in evaluation, with Scriven advocating the elevation of their construction to that of a *checklist theory*. Scriven described a *comlist* as the *criteria of merit* checklist. He detailed several advantages of checklists and the following have been identified as particularly relevant for evaluation of this study:

- Checklists … reduce errors of omission.
- Comlists reduce … the tendency to see what one wants to see in a mass of data … by forcing a separate judgement on each separate dimension and a conclusion based on these judgements.
- Checklists often incorporate huge amounts of specific knowledge about the particular evaluands for which they have been developed. (selected from Scriven, 2005, pp. 3-4)

Scriven described a *core comlist* as

> … defining the general notion of merit … (Scriven, 2005, p. 5)

He recommended that the lists should be

- complete and
- concise . (summarised from Scriven, 2005, pp. 4-5)

and that the checkpoints that constitute *comlists* should

- refer to criteria and not mere indicators
- be nonoverlapping

- be commensurable and
- be clear.
- confirmable. . (summarised from Scriven, 2005, pp. 4-5)

Scriven distinguished between the *criteria of merit* and indicators. Scriven's notions of *comlist* and *indicator* have been adapted for this study. Table 5.4 illustrates a *comlist* associated with the evaluands for this project. The criteria for each principal notion to be evaluated have been associated with indicators of achievement. Note that the usage of *indicator* in this sense is not necessarily commensurate with that of Scriven. (Scriven, 2005) However, the 'indicators of achievement' for each of the criteria have enabled increased specificity of the criteria to facilitate the 'checking' process. This has been illustrated in Table 5.5.

The researcher has recognised the usefulness of the checklists and has used their specificity to achieve a more comprehensive but systematic view of the complexities of the evaluation process.

Table 5.4: A general *comlist* for the objectives of this study

Criteria	Indicators of achievement
1. Student learning	
1. Assessment tasks target and reward the levels of knowledge and skills identified in the learning objectives	Correlation of deconstructed tasks, solutions and marking criteria using Bloom's Taxonomy
2. Students have demonstrated achievement of the discipline learning objectives	Assessment results/grades; Teacher responses/comments
3. Student have demonstrated confidence in their own achievement	Survey responses/comments
4. Students *know what the know* and *know what they don't know*	Correlation between rankings of topic assessment results and rankings of student topic confidence
2. Teaching	
1. Facets of the subject's presentation support student achievement of the learning objectives	Positive student survey responses on the value of each facet to their learning; Positive attendance and submission rates; Improvements in student marks and grade distributions; Supportive comments and responses from participating teachers and markers.

Source: Adapted from Scriven, 2005, pp. 4-8

Table 5.4 (continued): A general *comlist* for the objectives of this study

3. Assessment	
1. Assessment is aligned with the subject learning objectives and the student learning experiences	Positive student survey responses; Improved completion rates; Improved submission rates; Improvements in student and grade distributions; Comments and responses from participating teachers and markers. Correlation of deconstructed tasks, solutions and marking criteria using Bloom's Taxonomy
4. Alignment of teaching, learning and assessment	
1. Assessment is aligned with the subject learning objectives and the student learning experiences	Positive comments and responses from participating teachers and markers. Increased cognitive demand revealed by deconstruction of tasks, solutions and marking criteria using Bloom's Taxonomy
5. Statistical thinking	
1. Assessment is aligned with the subject learning objectives and the student learning experiences	Increased cognitive demand revealed by deconstruction of tasks, solutions and marking criteria using Bloom's Taxonomy Student marks and grades static or improved; Comments and responses from participating teachers and markers.
6. Critical and evaluative thinking	
1. Assessment is aligned with the subject learning objectives and the student learning experiences	Increased cognitive demand revealed by deconstruction of tasks, solutions and marking criteria using Bloom's Taxonomy Student marks and grades static or improved; Comments and responses from participating teachers and markers.

Source: Adapted from Scriven, 2005, pp. 4-8

Table 5.5: An example of an expansion (of checkpoint 1.1 of Table 5.4) of the general *comlist*

1. Student learning 1. Assessment tasks target and reward the levels of knowledge and skills identified in the learning objectives	
1. Assessment questions target and reward the levels of knowledge and skills identified in the learning objectives	Deconstructed questions using Bloom's Taxonomy
2. Assessment solutions target and reward the levels of knowledge and skills identified in the learning objectives	Deconstructed solutions using Bloom's Taxonomy
3. Assessment marks target and reward the levels of knowledge and skills identified in the learning objectives	Deconstructed marking criteria using Bloom's Taxonomy

5.2.4 An adapted evaluation model

This study has investigated

> … the contention that a pedagogy aligned through specified objectives enhances student learning in selected undergraduate university subjects. (See Chapter One of this thesis)

The notion of alignment is complex, and has been discussed in Chapter Four paragraph 4.1.5. Of relevance to consideration here is the fundamental idea that common objectives can focus the teaching, learning and assessment towards the desired learning outcomes. To evaluate how successfully the programs have been aligned, it has been necessary to examine the contention with regard to several foci:

- the objectives themselves as the agents of alignment;
- the subject presentation (incorporating all aspects of the pedagogy) and its importance to student learning;
- the subject assessment as an agency for alignment and as a vehicle for evaluating student learning;
- the subject evaluation; and
- student learning.

As student learning is presumed dependent upon all of the above aspects of the program, investigation of it has formed an important dimension of almost all of them.

In Paragraph 5.1.5, the modified Alexander and Hedberg model (Bain, 1999) (see Appendix 5.3) was identified for adaptation for this study. The resulting model has encompassed four phases: *observation* (detailed in Table 5.6a), *development* (detailed in Table 5.6b), *implementation* (detailed in Table 5.6c) and *review* (detailed in Table 5.6d).

A case study tracking statistical learning across five sessions at the University of Wollongong has been described in Chapter 6 of this thesis. The framework detailed in Table 5.6a is applicable to evaluation of the *observation* phase of that study. The frameworks described in Table 5.6b and Table 5.6c have been cyclically applied across the four sessions over which the teaching/learning framework was implemented. As these implementations formed part of an action research project, refinements were made to the teaching/learning framework as a result of the formative evaluation. The *review* phase was constructed to facilitate the summative evaluation of the study.

A single implementation of a teaching/learning framework constructed to promote critical and evaluative skills has been detailed in Chapter 7. For this study there was no *observation* phase, and the *development* phase preceded the teaching session.

The foci, purposes, evidence and methods are displayed in Tables 5.6a, 5.6b, 5.6c and 5.6d for the four phases.

Table 5.6a: An integrated evaluation framework for both case studies for the *Observation Phase*[*]

Focus	Purpose	Evidence and methods
Objectives	To examine the existing set of learning objectives: • to determine if they articulated the desired learning outcomes • discipline knowledge and skills; • university graduate attributes.	Researcher's annotated journal; Peer and expert review.
Presentation	To describe the teaching/learning framework with respect to its impact on student achievement of the learning objectives by: • identifying what *works*; • examining what *does not work*; • formulating/detecting ways of improving/remedying learning outcomes.	Researcher's annotated journal; Peer and expert review; Student surveys; Assessment marks and grades.
Assessment	To determine if existing assessment addresses the desired level of knowledge and skills, identifying successful and less successful strategies.	Student marks and grades; Deconstruction of the questions and responses using Bloom's taxonomy.
Evaluation	To determine if current evaluative practice allows broad evaluation of: • all aspects of the implemented curriculum; • student achievement of the learning objectives.	Researcher's annotated journal; Peer and expert review; Student surveys; Assessment marks and grades; Deconstruction of the questions and responses using Bloom's taxonomy.

(Adapted from Bain, 1999, p.168)

Table 5.6b: An integrated evaluation framework for both case studies for the *Development Phase*[*]

Focus	Purpose	Evidence and methods
Objectives	To reframe the existing set of objectives to: • articulate desired levels of discipline knowledge and skills; • include desired graduate attributes • promote alignment of teaching, learning and assessment.	Researcher's annotated journal; Peer and expert review.
Presentation	To develop a teaching/learning framework to promote achievement of the learning objectives by: • re-implementing what *worked*; • removing what *did not work*; • including new supportive innovations..	Literature review; Researcher's annotated journal; Peer and expert review; Student surveys; Assessment marks and grades.
Assessment	To design assessment that aligns with teaching and learning experiences and demonstrates achievement of the objectives.	Student marks and grades; Deconstruction of the questions and responses using Bloom's taxonomy.
Evaluation	To identify evidence sources for evaluation of • the implemented curriculum; • student learning 1. as assessed in the subject assessment, particularly the final exam; 2. as evidenced in student perceptions of their learning.	Researcher's annotated journal; Peer and expert review; Student surveys; Assessment marks and grades; Deconstruction of the questions and responses using Bloom's taxonomy.

(Adapted from Bain, 1999, p.168)

Table 5.6c: An integrated evaluation framework for both case studies for the Implementation Phase*

Focus	Purpose	Evidence and methods
Objectives	To determine if the objectives enable alignment of teaching, learning and assessment.	Researcher's annotated journal; Student surveys; Peer and expert review.
Presentation	To refine the teaching/learning framework with the intention of supporting student achievement of the learning objectives by • identifying what *works*; • reviewing what *does not work*; • improving learning outcomes.	Researcher's annotated journal; Peer and expert review; Student surveys; Assessment marks and grades.
Assessment	To determine if assessment aligns with the teaching and learning experiences and addresses the level of knowledge and skills declared in the subject learning objectives	Student marks and grades; Student surveys; Deconstruction of the questions and responses using Bloom's taxonomy.
Evaluation	To refine evaluative practice to promote evaluation of: • all aspects of the implemented curriculum; • student learning.	Researcher's annotated journal; Peer and expert review; Student surveys; Assessment marks and grades; Deconstruction of the questions and responses using Bloom's taxonomy.

(Adapted from Bain, 1999, p.168)

Table 5.6d: An integrated evaluation framework for both case studies for the *Review Phase*[*]

Focus	Purpose	Evidence and methods
Objectives	To examine the role of the objectives in: • construction of an aligned teaching/learning framework; • supporting student learning.	Researcher's annotated journal; Student surveys; Peer and expert review.
Presentation	To determine if the teaching/learning framework has supported student achievement of the learning outcomes.	Researcher's annotated journal; Peer and expert review; Student surveys; Assessment marks and grades.
Assessment	To determine if • if assessment has addressed the level of knowledge and skills declared in the subject learning objectives; • there is evidence of improvement in student learning.	Student marks and grades; Deconstruction of the questions and responses using Bloom's taxonomy.
Evaluation	To determine if evaluative practice allows evaluation of: • all aspects of the implemented curricula; • student learning.	Researcher's annotated journal; Peer and expert review; Student surveys; Assessment marks and grades; Deconstruction of the questions and responses using Bloom's taxonomy.

[*] The *Development* through *Review Phases* were cyclically applied across four implementations in STAT131.

(Adapted from Bain, 1999, p.168)

5.3 The teaching/learning framework: *sub-atomic* evidence

Reference to professional experience and extensive literature review helped to identify appropriate facets for incorporation in the teaching/learning framework (which is discussed in more detail on Chapter 4). Evaluation of the framework needed to encompass an examination of the impact of each facet upon student learning. Table 5.7 enumerates some detail here for cross reference with other aspects of the evaluative processes.

Table 5.7: Focus detail of the teaching/learning framework

Focus	STAT131	Introduction to Accounting Theories and Philosophies
Facilitating student learning – subject presentation	Experiential learning Defined objectives Collaborative learning Authentic tasks Technology Modelled thinking Personal workbooks Marking guides	Experiential learning Defined objectives Collaborative learning Authentic tasks Technology Modelled thinking Personal workbooks Marking guides
Constructing an aligned pedagogy	Revised taxonomy Behavioural objectives	Revised taxonomy Behavioural objectives
Specifically targeted skills	Statistical thinking	Critical and evaluative thinking

Chapter 6

Case Study: Doing the sums - improving a student's perspective of statistics

"I didn't know I was to have a party at all," said Alice; "but if there *is* to be one, I think *I* ought to be the one to invite the guests."

"We gave you the opportunity of doing it," the Red Queen remarked: "but I dare say you've not had many lessons in manners yet?"

"Manners are not taught in lessons," said Alice. "Lessons teach you to do sums, and things of that sort."

"Can you do Addition?" the White Queen asked. "What's one and one and one and one and one and one and one and one and one and one?"

"I don't know," said Alice. "I lost count."

"She ca'n't do Addition," the Red Queen interrupted. "Can you do Subtraction?"

"Nine from eight I ca'n't, you know," Alice replied very readily: "but-"

"She ca'n't do Subtraction," said the White Queen. "Can you do Division? Divide a loaf by a knife – what's the answer?"

"I suppose – " Alice was beginning, but the Red Queen answered for her. "Bread-and-butter, of course." …

"She ca'n't do sums a *bit!*" the Queens said together, with great emphasis.

From Queen Alice in
Through the Looking Glass by Lewis Carroll p. 221

6.0 Background

This case study tracks the learning of undergraduate statistics students at the University of Wollongong in New South Wales, Australia. These students are not atypical of other Australian students studying at other universities and indeed, not dissimilar to their global counterparts. Nevertheless, this paragraph provides background description of the study.

6.0.1 A need to engage more students in the study of statistics

Students in New South Wales differ in some respects from those in other Australian school systems in that they receive limited exposure to the subject of statistics in secondary schools. Neither has there been any requirement to support the study of statistics with technology (SSAI, 2005). The subject of statistics has consequently not generally been highly regarded by students and they have, more often than not, approached the study of statistics with a negative mindset:

> Statistics has a poor image and profile among students, parents and the general public. (SSAI, 2005, p. 9)

It has not been surprising then that students have traditionally found the study of statistics difficult (SSAI, 2005). Formulae are complex and unfamiliar, and statistics challenges students to solve problems differently to the traditional approaches they have adopted in solving mathematical problems. There may be no single answer, but multiple alternatives. Context does not cloud meaning but is pivotal to it. There is a focus on solving real world problems (Moore and Cobb, 2000; Cobb and Moore, 1997; Gal and Garfield, 1997). High school teaching has to date failed to address these aspects of the discipline effectively (SSAI, 2005). To date, before university studies, students have only been introduced to very basic procedures accompanied by undeveloped concepts, and provided with little or no modelling of statistical thinking (although changes to address these issues are in progress). Thus for most students, their enrolment in subjects such as this, has represented an excursion into unfamiliar territory, and they have generally approached it with attitudes ranging from disinterest to trepidation. The weak grounding of the curriculum has inspired a need for redress of the resultant *shallow learning*.

6.0.2 An aim for the study

The aim then has been to develop a teaching/learning framework that would engage students actively in learning to *think statistically* (Chance, 2002; Gal and Garfield, 1997). Such a framework should use exploration of data to develop understanding of the tools and procedures available to the statistician (Rossman, 1997; Moore and Cobb, 2000, Chance, 2002) and through the use of technology, should remove much of the drudgery of numerical calculation that can distract students from the reasoning processes that underpin statistical study (Rossman and Chance, 1999). Indeed Chance has claimed

> Aided by recent advancements in technology, “number crunching” no longer must dominate the landscape of the introductory course. Instead, we have the luxury of allowing our students to focus on the statistical process that precedes the calculations and the interpretation of the results of these calculations.” Chance, 2002, p. 1)

Construction of an effective learning environment is more often than not an evolutionary process resulting from spiralling action and reaction to student responses. The teacher’s reflective responses are grounded in experience and integrated from practice, the literature and peer discussion (Entwistle et al., 2000).

As a teacher, the researcher has drawn upon her years of classroom practice, collaboration with *experts* teaching in the discipline of statistics and her reading in attempting to construct a pedagogy that will effectively address desired learning. She has done this, however, recognising that student learning (or perhaps lack of learning) frequently drives changes in pedagogy and the consequent responses of students have inspired developmental changes. In a dynamic teaching environment, responses are made to the changing needs of its students and these in turn drive continuous refinements to the design of the teaching/learning framework.

6.0.3 A methodology

This study has tracked the progress of a pedagogical design aimed at developing *statistical thinking* in undergraduate students. Action research afforded the methodological approach that allowed documentation of the pattern of change across a *spiral* of five implementations of the teaching/learning framework (see Figure 2.3, Chapter 2) (O'Leary, 2005; Cresswell, 2002, Robson, 2002). Table 6.1 details the sessions encompassed by this study.

Table 6.1: Implementation cycles of STAT131 included in this study

Session	Action
Autumn, 2003	Observation→Evaluation→Design/refinement
	↓
Autumn, 2004	Implementation→Evaluation→Refinement
	↓
Spring, 2004	Implementation→Evaluation→Refinement
	↓
Autumn, 2005	Implementation→Evaluation→Refinement
	↓
Spring, 2005	Implementation→Evaluation→Refinement
	↓ Summative evaluation

Case studies have been questioned as research methodologies, but have become regarded rather as qualitative research strategies that might incorporate various methodologies (Mertens, 2005; O'Leary, 2005). The study covers five teaching sessions, and hence five largely independent cohorts of students. A small number of failing students from each session re-enrolled in the subsequent sessions but data on these individuals has not been available within the data collection structure. Hence for the purposes of this study the cohorts have been regarded as independent. Despite the breadth of the study, this research has been specifically included as a 'case study' because the overarching theme has been the design, implementation and evaluation of *a teaching/learning framework* contextualised in *a specific undergraduate statistics subject* (O'Leary, 2005; Stake, 1995). The study also exhibits elements of other research methodologies (as described in Table 2.3 and discussed in detail in that chapter).

Data have been collected from multiple sources: student surveys, assessment data and the researcher's annotated journal diarising the critical features of each session. Appendix 6.1 through to Appendix 6.18 provide summaries of changes to the key facets of the teaching/learning framework for this subject, related student survey responses and assessment results across all of the sessions.

6.1 Descriptive detail

This case study examines the pedagogical design for a specific student body studying within a bounded context of a foundation course in statistics at the University of Wollongong in Australia. The following detail provides context in order to situate the needs of the students and the teaching environment in which the research has been undertaken. The study has progressed with the total support of the school, in particular the subject teachers and the cooperation of the students. Without such positive disposition and involvement it would not have been possible for the researcher to observe existing practice, and to consequently design, implement and evaluate strategies targeted at improving the depth of student learning by leading students to *think like the expert.*

6.1.1 A time and place …

The study has taken place across five teaching sessions between 2003 and 2005 at the University of Wollongong which is situated south of Sydney (Australia) near the New South Wales south coast. It is a highly regarded institution (listed twice in the Times top 200, and the winner of the inaugural 2006 Commonwealth University of the Year Award for Community Engagement from *The Times Higher Education Supplement* and the Commonwealth Association of Universities) and draws its students both from Australia and overseas. It has a student enrolment of approximately 23,000 students of which about 9,000 are international students. (University of Wollongong, n.d. (a))

At the University of Wollongong, the School Mathematics and Applied Statistics is staffed (in Statistics) by teachers who are also active researchers and consultants. These staff members are well situated to impart not only the discipline knowledge and skills, but also to *model* the *expert thinking* that is essential for their application to problems in a *real life* context.

6.1.2 Working with experts

STAT131 had been designed and coordinated (autumn session – session 1) by a teacher committed to addressing statistical learning in an environment that engages students on all levels. The coordinator had received recognition nationally and internationally for her work in statistical education. Consequently, this subject has had a history of innovative teaching practice, with learning promoted through experiential learning, authentic tasks and the use of technology. Evaluation practice had encompassed data both from assessment and student surveys.

In the spring sessions (session 2), the coordinator's role for STAT131 was handed to another teacher. This coordinator was also committed to student learning. He was supportive of the existing curriculum, and did not implement drastic changes to it. Spring sessions were invariably modelled on the preceding autumn sessions.

Tutors varied in their level of expertise. Some were experienced academics, some doctoral students in statistics or mathematics education, while others were discipline-competent undergraduates. From 2005, as a consequence of Occupational Health and Safety Regulations, all tutors were required to complete a basic course in order to tutor at the university.

6.1.3 Working with novices

There were between 150 and 200 students enrolled each session (Table 6.4). There was little reason to suspect substantial differences in the patterns of enrolment between the corresponding sessions, except that a small number of unsuccessful students from the autumn session re-enrolled in both the spring and autumn sessions. Across all sessions, the students were predominantly male (approximately 80%), and 20-30% were international students (see Table 6.4).

6.1.4 Content and skills

The teacher has taught mathematics and statistics at university for over ten years. As a committed teacher, she has sought not only to instil in her students the requisite knowledge and skills, but also to inspire an understanding of *statistical* and *mathematical thinking* that generates an appreciation of the worth of the subject. The subject that has formed the basis for this study has serviced several courses at the University of Wollongong, and has covered fundamental statistical concepts and procedures, with

> The focus … on understanding concepts and evidenced-based decision making, and topics (that) covered exploratory data analysis, probability models, regression and hypothesis testing. (Morris, Porter and Griffiths, 2007a, p.1)

The mechanical tedium of formulae selection and numerical calculation was relieved by use of the statistical software package SPSS.

STAT131 has been a compulsory subject for most of its students, and the majority of these were drawn from degree programs in Information Technology and Computer Science. Many of the students were in the first year of their degree program.

6.2 An existing infrastructure

At the start of the study, the subject designer had already instituted a specifically tailored pedagogy centred on practical engagement of students. She aimed to actively engage students in *thinking statistically* and had endeavoured to create a learning environment that supported this end. Key features of the existing infrastructure in session 1, 2003 are described here and should be read with reference to the evaluative framework given in Table 5.2a.

6.2.1 Lectures

Lectures were well structured and students had access to the PowerPoint slides associated with them. The lectures targeted multiple learning styles including active, reflective, theoretical and pragmatic and were not confined to more narrow perspectives such as auditory or visual perceptions (Watkins et al., 2002). (See paragraph 3.3 for further discussion of learning styles.) Lectures were interactive and although attendance was not compulsory, as they were aligned with the exercises in the laboratory classes, it was encouraged.

6.2.2 Laboratory classes

In laboratory classes students worked independently through tasks contained in a laboratory manual. They built up a portfolio of solutions to the tasks and supplementary 'pen and paper' questions in Word documents and these were checked for completion by a tutor. The tutor's role varied between classes, with some electing to support student learning by review of relevant theory, while others facilitated independent completion of the tasks. Many students viewed the tutor's role as support in their use of the unfamiliar software (SPSS).

Tasks in the laboratory manual were designed to ground the theoretical aspects of the lecture material, and involved guided application of the theory to solutions of contextualised problems involving real data. The manual contained copies of the PowerPoint lecture slides and some reference material for SPSS.

6.2.3 Assessment

Assessment of learning included:

- three assignments based upon: exploratory data analysis; probability models; and evidence based decision making;
- a midterm quiz based on pen and paper type questions similar to those in final examination. In 2003 marks were also allocated to student prepared summary notes;
- a presentation based on assignment responses;
- a laboratory mark that recognized completion of the set tasks and
- a final exam.

Feedback on continuous assessment was prompt, usually a week from attempt/submission. No feedback was given on the final exam.

Assignments offered opportunities for collaborative interaction with a partner (see Chapter 4, paragraph 4.3.1 and this chapter, paragraph 6.3.3). Students discussed the tasks with their partner but most tasks required responses to different data for each partner. There were also questions that were common to both partners and students were required to submit independent responses to these questions.

The assignments were marked using a prescribed marking scale, but there was some latitude for the marker in allocating marks within this scale. Student feedback was achieved through annotation of the assignment and access to the complete solutions once the assignments had been returned.

Some questions targeted *deeper learning.* However, most questions elicited lower order cognitive responses and marks were allocated in many cases for calculation and repetition of fact or process.

6.2.4 Evaluation

At the time of the researcher's engagement in this study, the subject afforded an opportunity to work in a teaching environment that allowed insight into professional expertise grounded in evidence-based practice. Refinements made to the pedagogy reflected responses to identified needs. It was a teaching/learning environment that was receptive both to innovation and the enthusiastic enterprise of this less experienced researcher.

The subject teacher's response to her classroom was based upon student assessment data, an online student survey of her teaching practice, peer collaboration and reflective practice. Her existing evaluation strategies explored student responses to the subject presentation and included opportunity to comment on their learning. Such practice afforded strong groundwork for *this* teacher/researcher in her attempts to effectively explore the learning of students in this case study.

6.2.5 Learning to tread in expert footsteps

The pedagogy included opportunities for students to experientially engage in

- collecting data
- exploring and analysing that data by applying theoretical concepts
- using technology to support the action
- resolving *authentic* problems.

The learning tasks were designed to promote the *statistical thinking* of the expert and to challenge student beliefs and practices. In autumn 2003, students responded to a survey on the Australian commitment of troops in Iraq. In the multicultural social context of the classroom, the responses to this exercise proved enlightening to both the students and this teacher/researcher!

6.3 Promoting learning: enhancing the framework

Anticipated learning in this subject included:

- development of evaluative and analytical reasoning, in particular, *statistical thinking;*
- the acquisition of the prescribed concepts and skills;
- the promotion of *meta-cognition*;
- fostering teamwork, an appreciation of social and cultural diversity and the ability to function in a global environment; and
- the use of technology in evidence-based decision making. (University of Wollongong, n.d. (b))

The teacher's aim was to institute teaching/learning strategies that facilitated such learning. However, enhancing a teaching/learning framework is a developmental process and hence ongoing throughout the study. Each implementation highlighted *what worked* and *what did not work*, and this drove the subsequent implementation. Reading of the literature and collaboration with peers also propelled changes and it is for this reason that the design and development phases have been treated simultaneously in this paragraph. Table 5.2b and Table 5.2c specify the evaluative frameworks that have provided structure to these phases.

With solid foundations already in place, the teacher set out to construct an enhanced pedagogy. Fundamental to her design was the *alignment* of all facets of teaching, learning and assessment through explicitly and behaviourally defined learning objectives (Anderson and Krathwohl, 2001; Angelo, 1999; Biggs, 1999).

But teaching is not just about providing clear focus on desired learning outcomes. Engagement of students in effective learning activity also requires the careful orchestration of appropriate tools and tasks to support them. Enhancing the teaching/learning framework for this unit has required an eclectic approach incorporating reference to observation and experience, discussion with peers and reference to the abundant literature. Chapters 3 and 4 provide more detailed discussion on identified strategies than has been included here. Whereas Figure 3.3 depicts the interrelationships between more general aspects of relevant learning theories in the devised teaching/learning framework, in Figure 6.1 some of the more theoretical inputs have been replaced by their strategic counterparts to give a practical perspective of the implemented framework for this subject.

Please see print copy for Figure 6.1

Figure 6.1: Teaching/learning framework
showing implemented strategy inputs.
Adapted from Morris, Porter and Griffiths, 2004a, p. 98

6.3.1 Defined learning outcomes

Remnants of high school teacher training still underpin the teacher's approaches to her classroom teaching. Global specification of teaching aims and objectives have remained the basis of her lesson design. In the shadow of the influence of Bloom (1974), she has attempted to define her students' expected learning behaviourally, and her vigilance in detecting the signs (or lack thereof) of student achievement of these objectives has inspired revision, reteaching or changes to her teaching tempo.

The advent of outcome-based curriculum to years eleven and twelve in the high school system in N.S.W. (with the first consequent Higher School Certificate examination in 2001) (Board of Studies, 2001) re-affirmed the focus of high school teaching on addressing outcomes, and assessing and reporting against those outcomes.

In her experience of its implementation, the teacher recalled an initial inertia to its acceptance. To the teacher, teaching both in high school and at university at the time, there seemed to be little difference between her existing high school practice and that which an outcome based model might require. However, closer examination highlighted some advantages.

The outcome-based model reinforced the necessity to assess and mark against outcome achievement, and the consequent *alignment* promoted a greater transparency in the teaching, learning and assessment loop. In addition, at the senior levels, there was a requirement for prior advertisement of the outcomes to be assessed. This presented opportunities for encouraging students to develop greater awareness of what learning they needed to demonstrate and hence to prepare with greater focus. To the teacher, there seemed no reason to suppose that such a model appeared would not offer similar advantages at the university level. This insight presented the beginnings of the inspiration for researcher to undertake this study.

The teacher has used behaviourally framed objectives to *align* aspects of the teaching learning frameworks (Biggs, 1999). Alignment has also been used to prevent cognitive mismatch between teaching and learning (Lugenbehl, 2003; Biggs, 1999). The revision of Bloom's taxonomy (Anderson and Krathwohl, 2001) has facilitated the comparison of *teaching intent* with *assessment practice*. Thus all learning tasks, instructional activities, assessment contain explicitly stated objectives and marking criteria and feedback for assessment relate directly to them (Anderson and Krathwohl, 2001).

6.3.2 Marking guides as agents of scaffolding and feedback

In this study, the marking criteria have fulfilled multiple purposes:

1. Alignment with the subject learning objectives:
   Biggs (1999) advocated *constructive alignment* of the teaching/learning/assessment framework. The development of *scaffolding* (see 4.1.4) and the *marking criteria* (see 4.3.6) is discussed in Chapter 4. For this framework, the specified learning objectives for each assessment task were deconstructed to build the marking criteria. Hence these criteria described fundamental behavioural indicators of achievement of the learning objectives (including discipline and generic knowledge and skills

enabling their transfer to other contexts) (James et al., 2002; Sadler, 1998), and were checked as achieved/not achieved according to student responses. Table 4.5 provides an exemplar of such a set of criteria.

2. Scaffolding for supporting student learning:
   Copies of the explicit criteria (divorced of marks) were given to students with assessment tasks as 'marking guides'. The guides served as scaffolding for student responses (Ausubel, 1978; Sadler, 1998). Table 4.2 provides an example.
3. Detailed feedback:
   As the checked criteria identified what learning had been demonstrated or had failed to be demonstrated, students had direct (and timely) feedback that was both task and standards related (Biggs, 1999) and effected positive motivational feedback (Black and Wiliam, 1998). It also highlighted knowledge and skills in need of remediation. Used in conjunction with the task solutions and modelled exercises in the laboratory manual, the checked criteria provided a remediation path for addressing their learning deficits (Sadler, 1998; Black, 1998; Crooks, 1988).
4. Transparent marking
   Marks were allocated on the basis of 1 for achieved/demonstrated and 0 for no achievement/demonstration and this provided a mechanism for minimising subjective allocation of marks (see Chapter 4, paragraph 4.3.6). For this reason, in addition to student access to the criteria before submission and their explicit nature, students perceived assessment as fair (James et al., 2002).

Biggs (1999) did however acknowledge the potential for *misalignment* to promote *shallow learning* even if specific learning objectives have been defined. This can result from reporting assessment outcomes against a bell curve rather than the defined outcomes (see Chapter 4, paragraphs 4.2.1 – 4.2.3 for a more elaborate discussion).

6.3.3 Collaborative Learning

The researcher has acknowledged her adherence to constructivist learning philosophies (see Chapter 3) and hence has targeted learning strategies in keeping with those philosophies. Such philosophies support active engagement in a socially interactive learning environment and appropriate contexts for learning tasks. These strategies underpin collaborative learning (Hernandez, 2002; Watkins et al., 2002; DeMulder and Eby, 1999). Benefits deriving from collaborative learning include:

- improvement in individual performance (Pfaff and Huddlestone, 2003; Livingstone and Lynch, 2000; Johnson et al., 1998);
- promotion of persistence, transferability of knowledge, social skills and intrinsic motivation (Johnson et al., 1998);
- improved management skills, communication skills and multiethnic relations (Watkins et al., 2002);
- promotion of higher order thinking and *deeper learning* (Hernandez, 2002; Biggs, 1999);
- the viewing of problems from multiple perspectives (Hernandez, 2002, DeMulder and Eby, 1999);
- the potential for encouraging a critical reflection that impacts upon students' value systems and senses of social responsibility (Livingstone and Lynch, 2000);
- opportunities for critical evaluation, judgement, mentoring and advising. (DeMulder and Ebe, 1999)
- a foundation upon common goals rather than a promotion of competition (Johnson et al., 1998); and
- the potential to encourage critical reflection that impacts upon students' value systems and senses of social responsibility (Livingstone and Lynch, 2000).

In this study, collaborative learning has been affirmed in classroom practice and has been continually enhanced in the continuous assessment structure throughout this study. The notion of parallel questions (same questions applied to different data sets) has been supplemented by complementary questions. These questions require use of a team partner's results to supplement an individual's response. This type of question serves to:

- promote interchange of ideas and understanding;
- foster active discussion of statistical knowledge and skills, enhancing their ability to *think* statistically;

- identify errors in an individual's or their partner's understanding before submission and thus the potential to rectify misconceptions without penalty; and hence
- confirm the responsibility for learning in the hands of the learner.

The focus of student presentations also changed across the sessions. It shifted from addressing the generic skills of public presentation, to the active demonstration of *statistical thinking* in an *authentic* context.

6.3.4 Experiential Learning

Deeper learning is promoted by students actively engaging students on problem-based learning tasks that demand higher order cognitive thinking (Biggs, 1999). Such tasks should offer opportunities for students to cyclically act, reflect on the action, absorb new learning and apply that learning in new contexts. Students may need to repeat steps to achieve the desired learning (Kolb, 1984). Watkins et al. (2002) extended Kolb's learning cycle to include an overarching cycle that initiates *learning to learn* (see Figure 3.2). *Meta-cognition* is discussed more fully in Chapter 4, paragraph 4.1.1).

6.3.5 Authentic Tasks

Authentic tasks address learning of discipline knowledge and skills in relevant and meaningful real life contexts. They enable students to construct meaning by engaging them in *thinking like the expert* (James et al., 2002; Montgomery, 2002). Data have been collected from the students themselves and used in analysis. Other real data sets have resulted from more global statistical collections.

6.3.6 The laboratory manual as a portfolio of learning

A learning portfolio is generally regarded as a collection of a student's work that reflects their learning and generally incorporates some elements of critical reflection evidencing *meta-cognitive* development (Klenowski, 2006; Biggs, 1998). For benefits to be maximised, the student needs to *own* the included work (Klenowski et al., 2006; Liu et al., 2004; McElwee et al., 2002).

The challenge in this study was to design a student workbook that incorporated:

- guided tasks that modelled *expert statistical thinking* and addressed the requisite discipline knowledge and skills;
- opportunities for students to transfer the modelled *thinking* to new contexts;
- structures designed to reinforce *meta-cognition*; and
- strategies that would promote ownership of the learning displayed.

While undertaking this study, the researcher worked on the case study discussed in Chapter 7. The main teaching intent in that subject was to expose students to critical and evaluative thinking in an accounting discipline. In identifying appropriate strategies for fostering the required higher order cognitive skills, the researcher encountered recent research attesting to the successful use of student portfolios (See paragraphs 4.3.8 and 7.3.6). As her reading explored the concept more fully and in the light of classroom practice, the laboratory manual of the current study took on many of the features of the learning portfolio.

Midway through the study, 'learning framework' tasks were included to facilitate organisation of complex detail and concepts, and some opportunities for reflecting on learning were added. To inspire focus on completion of the tasks, students were permitted to take the manual into the final exam. This motivated students to ensure that it was complete in order to derive maximum support.

However, across several sessions, students acquired previous sessions'/other students' solutions to the tasks and 'ownership' lost its impact. Many students had not actively constructed the learning, but 'borrowed' this learning. So the effectiveness of the manual as a learning tool was depleted. A more efficient method of addressing this dilemma has been instituted since 2005. Students subsequently attempted quizzes based on the learning tasks and this has revitalised student endeavours to 'own' the work in their manuals in order to demonstrate the required learning in each quiz.

The manual-to-portfolio transformation is still an evolutionary process in the work of the teacher/researcher. Promotion of independent and reflective learning is the target that inspires the ongoing effort.

6.4 A subject in parallel

Throughout this study, the teacher/researcher was also teaching in a second subject, STAT151/252. This subject primarily targeted students of the biological sciences and most students had already progressed beyond the first year of their undergraduate studies. The teaching content, experiential engagement of students, assessment regimen and the learning objectives were very similar to STAT131, and it was only offered in the second session of each academic year. The two subjects exhibited many differences (including gender distribution, non first year students etc.) and there were different teaching constraints, however, both subjects addressed the teaching/learning of statistics in a data-driven environment. It is the similarities between the two subjects that have prompted the teacher to implement comparable teaching/learning frameworks and the researcher has published comparative results on her work. (Morris et al., 2004b).

In 2003, the researcher piloted the concept of alignment through the subject assessment in STAT151/252. The use of *marking guides* in facilitating alignment between *intent* and *practice* in assessment and for focusing student responses on expected learning were targeted for testing. The process highlighted many shortcomings, and enabled address of these before the implementation in the next cycle of STAT131. Throughout the study this subject continued to provide a solid ‘testing ground’ for the implementations in STAT131, and in turn changes implemented in that subject impacted positively upon teaching learning in STAT151/252.

6.5 Evaluation strategy

In Chapter 5 the dilemmas faced in evaluating teaching innovations are detailed and the general *raison d’etre* for the evaluative practices adopted for both case studies within this work are provided. The strategies employed by the researcher in this study were modelled on the less formalised practices of the teacher and the foundation for their systematisation was in a framework already in place at the time of her engagement.

The process of evaluation was ongoing, modelling the practice of the committed and reflective teacher. It tracked and documented existing effective practice in the *Observation Phase*, sought best practice during the *Design Phase*, enacted the plan in the *Implementation Phases*, and informed further teaching through the *Evaluation*

Phases. Descriptions of the evaluation framework are given in Tables 5.6a – 5.6e. The *evaluand* was generally defined in Chapter 5 (paragraphs 5.2.1 – 5.2.2) for the cases studies of this thesis. The ensuing paragraphs provide greater detail for the *evaluand* for this case study, the evidentiary sources and to facilitate application of the evaluation framework.

6.5.1 Describing the *evaluand*

Effective evaluation requires an accurate picture of exactly what is to be evaluated. In a constructivist classroom concepts such as ‘student learning’ and ‘teaching/learning framework’ are complex and difficult to describe precisely and hence evaluation may be imprecise. Table 6.2 summarises the focus detail for examining the impact of teaching/learning upon student statistical thinking for this case study. In it the researcher has identified the following foci as areas for investigation:

- facilitating student learning – subject presentation;
- constructing an *aligned* pedagogy that addresses and checks *intent* and *practice*; and
- specifically targeted skills – *statistical thinking*.

Table 6.2: Focusing the evaluation of the learning framework for STAT131*

Focus questions		Focus detail	
What is to be evaluated?	The impact of all facets of the teaching/learning framework generally on student learning as defined by the objectives and more specifically on statistical thinking.	Facilitating student learning – subject presentation	Experiential learning Defined objectives Collaborative learning Authentic tasks Technology Modelled thinking Personal workbooks Marking guides
		Constructing an aligned pedagogy	Revised taxonomy Behavioural objectives
		Specifically targeted skills	Statistical thinking

*Summarised from the detail in Tables 5.2 and 5.7

This evaluation has been made with reference to the evidentiary sources given in Tables 5.6a-e:

- researcher's annotated journal;
- peer and expert review;
- student surveys;
- assessment marks and grades; and
- deconstruction of the questions and responses using the revised Bloom's taxonomy (Anderson and Krathwohl, 2001).

Student learning has been evidenced in the students' results in the *aligned* assessment and their confidence in their learning, as reported in the student surveys. The impact of the teaching/learning framework upon learning has been evaluated with reference to:

- students' responses in the student survey;
- *experts*' comments in surveys (marker and tutor surveys);
- discussion and peer review of presentations and publications reporting on the case study as it progressed.

The alignments have been achieved through the specification of behaviourally framed objectives using the revised taxonomy of Bloom (Anderson and Krathwohl, 2001). Targeted skills have been checked through application of the revised taxonomy of Bloom in assessment questions, solutions and student achievement.

6.5.2 Seeking evidence: the annotated journal

The researcher's journal has documented her observations, experiences and reflections for the case studies included in this work. Annotations have been made by the discipline experts and associated teaching staff coordinating the teacher's subjects. These annotations include queries of the recorded detail, reflective comment and inclusions perceived as warranted. Hence the journal affords some record of discussion between the researcher and the primary teachers.

6.5.3 Seeking evidence: peer review

Throughout the term of this study, the researcher has presented her work to other professionals (Morris et al., 2006a; Morris et al., 2005a; Morris et al., 2005b; Morris et al., 2004c; Morris et al., 2004d) and has published some of her research findings (Morris et al., 2007a; Morris et al., 2007b; Morris et al., 2004a; Morris et al., 2004b). This has afforded many opportunities to discuss her ideas and practices. Feedback from her colleagues has proved invaluable in her determination of an appropriate teaching/learning framework.

6.5.4 Seeking evidence: student surveys

In assessing the success of the teaching/learning framework, the researcher has built upon the existing evaluation structures observed in 2003. As the researcher sought to augment her understanding of student learning, the student survey (an example has been included in Appendix 6.25) was continually refined across the five sessions. It gathered information on:

- attendance patterns;
- overall learning;
- perceptions of the importance of aspects of the teaching/learning framework to student learning;
- their perceptions of their own topic learning; and
- their perception of their achievement of the specified graduate attributes.

The survey's evolution across all implementations has been documented in Appendix 6.18. Appendices 6.1 to 6.16 provide summary detail based upon these responses to questions related to the facets of the subjects presentation, while Appendices 6.19 and 6.20 examine perceptions of topic learning and progress toward achievement of the graduate attributes. It should be noted that the number of facets of the subject's presentation increased greatly across the implementations (from five facets in 2003, to fourteen in spring 2004 and to sixteen in 2005). Nevertheless, there remained some measure of comparability across all implementations.

Students responded to the surveys online. Hence, whilst encouraging teaching staff to promote its completion, the researcher was limited in her ability to influence response rates. The response rates varied across the implementations, with spring 2005 achieving the lowest rate (13%) (see Table 6.3). For this reason alone, the data retrieved from this survey should be treated as suspect in terms of being representative of the cohort's views for that session. The low response rate in spring 2005 reflected a university wide trend, later tentatively identified as a student response to over surveying Melano (2006).

Table 6.3: Student survey responses for each implementation

Session	Number of Responses	Number Enrolled	Response rate (%)
Observation phase: Autumn 2003	63	159	40
Autumn 2004	101	192	53
Spring 2004	63	173	36
Autumn 2005	62	152	41
Spring 2005	21	160	13

Source: Student survey data

Although gender and funding status (domestic/international) were not surveyed in the student surveys prior to 2005, comparisons across all sessions (from university funding data) have been included here to provide a description of the relative stability of the profile of enrolled students in terms of gender and funding status (domestic/international) across the period of the study. Proportions of gender and funding status demonstrate some differences across the period, but such differences represent small numbers in the associated cohorts.

Table 6.4: Student enrolment for the observation phase and each implementation

Session	N[1]	Percentage Male		Percentage domestic students	
		Survey data[2]	Funding data[3]	Survey data[2]	Funding data[3]
Autumn 2003 (Observation phase)	159	[4]	76	[4]	90
Autumn 2004	192	[4]	81	[4]	70
Spring 2004	173	[4]	85	[4]	78
Autumn 2005	152	71	80	71	79
Spring 2005	160	91	89	67	76

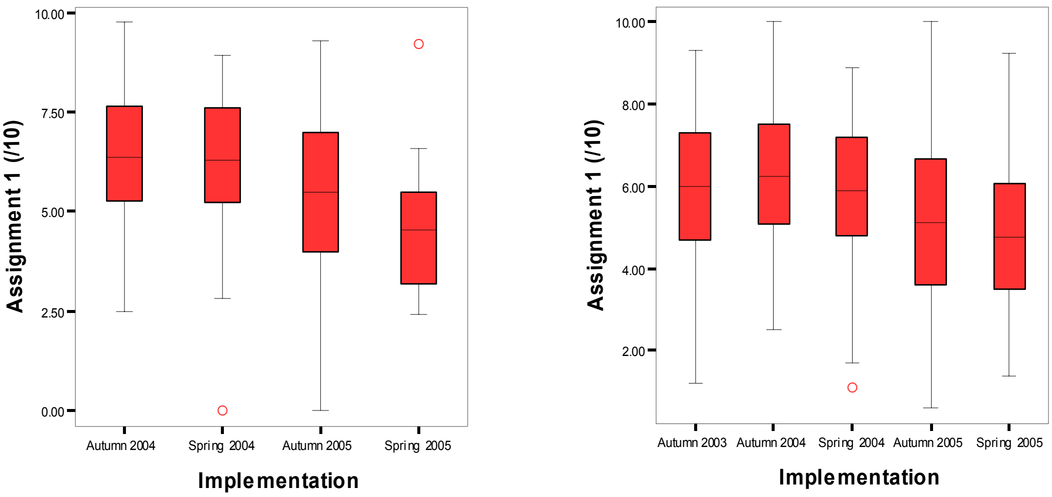
[1]Source: Student assessment data

[2]Source: Student survey data

[3]Source: University records of local students enrolled by funding classification (domestic/international)

[4]Not surveyed

Since student learning was of prime importance, reported student assessment marks (see Appendix 6.25, questions 27- 30) were compared to assessment data to detect bias that might be evident in the mark distributions. In Figure 6.2 below, the box plots of the distributions of the reported and actual marks for the first assessment task have been compared. The distributions of the reported marks did not differ greatly in terms of their centres, spreads or shapes from those of each session's marks extracted from the subject assessment data. The obvious exception to this was the final session, spring 2005. Although the mean for that session was not significantly different from the others, the distribution of the reported marks indicates that the survey data has not captured students from either extreme of the session marks. Hence the spring 2005 responses do not reflect students who fall into these levels of achievement.



Session	N	Mean	SE
Autumn, 2003	63	n/a	n/a
Autumn, 2004	101	6.5	0.16
Spring, 2004	63	6.3	0.18
Autumn, 2005	62	5.6	0.27
Spring, 2005	21	4.6	0.36

Session	N	Mean	SE
Autumn, 2003	157	6.1	0.14
Autumn, 2004	200	6.3	0.12
Spring, 2004	151	5.8	0.13
Autumn, 2005	159	5.3	0.16
Spring, 2005	136	4.9	0.15

Notes: Assignment marks were not reported in autumn 2003 student surveys (Observation phase)

Outliers are marked with the letter 'o'.

Figure 6.2: Comparison of reported Assignment 1 marks (scaled out of ten) with recorded assessment data across all implementations

6.5.5 Seeking evidence: student assessment

For an appropriately *aligned* teaching/assessment regimen, assessment should afford evidence of student learning. There are therefore two aspects to evaluation that need to be considered:

1. the *alignment*; and
2. the *marks*.

For the marks to be indicative of the intended learning, these too should be aligned with the intent of the assessment tasks. The revised taxonomy (Anderson and Krathwohl, 2001) was used to align the assessment intent with the marked responses of the students (See Chapter 3, paragraph 3.2.1). This task proved more difficult than had been anticipated. Although questions may target particular levels of knowledge and skills commensurate with the teaching/learning objectives, the researcher discovered that teachers may reward for lower order knowledge and skills (through allocating marks for their demonstration). Hence analysis of the questions in terms of the taxonomy should be matched with a parallel analysis of the acceptable responses. This difficulty arises more commonly within the fields of statistics and mathematics, where *marks* may be allocated for steps indicating low order knowledge and cognitive processing. Hence, *alignment* considerations have proved more complex than the researcher anticipated.

Marks then are only comparable if the cognitive demand has also been comparable. The assessment marks for all assessment for this subject have been referred to as the *student assessment data.* These marks were maintained by the subject coordinator and stored within the university system. The researcher has only accessed data for this study that has been stripped of all identifiers, even though the teacher has viewed the complete files.

The assessment regimen did vary slightly across the sessions, and the changes are evident in Appendix 6.21. However marks for corresponding assessment tasks have been scaled out of ten for the purposes of comparison. Common components of student assessment across all implementations were:

- assignment 1: basic data exploration and correlation between variables;
- assignment 2: model fit;
- assignment 3: confidence intervals and hypothesis tests;

- a midterm test;
- a laboratory tasks; and
- a final exam.

The final exam was worth 50 percent of each student's final mark.

All of the continual assessment was designed to be formative, with detailed marking criteria aligned with the assessment task objectives and complete solutions provided to the students.

6.6 Tracking the evidence: the teaching/learning framework

The students' perspective of the effectiveness of the teaching/learning framework was examined through responses to the student surveys. Each facet of the subject's presentation is discussed here, and the detail of students' responses and relevant demographic data have been summarised in Appendices 6.1 - 6.17.

Students were offered a range of responses reflecting their perception of the importance of each facet of the subject's presentation to their learning:

a. Not important
b. Of little importance
c. Moderately important
d. Extremely important.

For summary purposes, the last two options were aggregated and re-labelled as 'important'. Appendix 6.17 details the facets surveyed across all implementations ranked by the proportion of students perceiving them as important to their learning.

At the beginning of this study, the subject designer had introduced changes in the laboratory classes by including more authentic tasks that engaged students in the data gathering process. Prior to these innovations (see Figure 6.3 for comparative rankings of importance in 2002), the vast majority of students (95.7%) had rated *lectures* as important to their learning. However post 2002, the *laboratory classes* and the *solutions* to exercises and assignments assumed the higher ratings, each with proportions in excess of 90% (see Appendix 6.17).

Please see print copy for Figure 6.3

[1]Source: Evaluation files of Anne Porter, *Evaluation2002.*

Figure 6.3: Ranked order of the reported proportions of students perceiving surveyed facets as important to their learning. (2002)[1]

It has become evident throughout this study that all aspects of the teaching/learning framework are very closely related. Designing for alignment has ensured this. Despite the efforts of the researcher to address each aspect of the framework independently, the task has proved extremely difficult. Although alignment has promoted focused teaching, learning and assessment, the limited value placed upon the agents of alignment in students' survey responses (the *objectives* and the *marking guides*) might seem a little surprising! Students have commented on the connections between assessment and classroom learning in their open responses but these comments highlighted an apparent failure to notice the mechanism. The following comments from a committed student who worked continually throughout the session were elicited in an online interview and epitomise this perspective. As the teacher sought strategies that would inspire the sort of effort that this student had demonstrated, she was interested in analysing such responses and surveyed 'targeted' students to document their perceptions, in particular the perception of an *aligned pedagogy*:

> Question: Did you notice the 'objectives' listed at the start of each assignment?
> Response: Yes.
> Question: Why do you think they were placed there?

> Response: So that a student would know what subject matter was expected knowledge for the completion of the assignment.
> Question: Did you refer to them?
> Response: No.
> Question: Did you learn much from the assignments?
> Response: The assignments were really only small tutorial tasks. In fact, they took less time than the tutorial assignments each week. The majority of what was attempted in the assignments was attempted in labs, and as such, it was more of a confirmation of what I had learnt, and what would be expected of me in an exam (Targeted STAT131 student's response to an online survey, autumn 2004).

6.6.1 The impact of *lectures* on student learning (see Appendix 6.1)

There is a substantive difference in the proportion of students valuing *lectures* between the autumn (73-79%) and spring sessions (56-57%). This corresponded with changes in the lecturers. The subject designer was the subject coordinator and lecturer in the autumn sessions and the coordinator and lecturer in spring sessions modelled lectures and assessment very closely on autumn 2004.

Reported attendance also followed a similar pattern. The low proportions claiming to have attended lectures for more than ten weeks of the thirteen week session (14-19% for the spring sessions compared with 44-48% for the autumn session) were highlighted by the spring lecturer in discussions with the researcher. The more interactive style of the subject designer appears to have engaged more students and this was commented on by the spring lecturer in response to an online survey question from the researcher:

> Question: What ideas/ strategies do you have in mind that you would like to implement for improving teaching and learning in this subject?
> Response: … Add more activities in lectures. (Spring coordinator's response to an online survey, spring 2004)

6.6.2 The impact of *laboratory classes* on student learning (see Appendix 6.2)

The increased focus on learning in the *laboratory classes* from autumn 2003 has been mirrored in the students' perception of their value to their learning (between 82-91% regarding them as important), although survey results spring 2004 appeared to indicate a poorer perception of their value (76%, a ranking of 8 from a survey of 16 facets). Collaborative learning was encouraged as students actively addressed the tasks. There was some teacher instruction, but as this was not prescribed by the subject coordinator, student experiences differed from class to class. Most students required support with use of SPSS for analysis and this challenge formed the *laboratory class* focus for many students.

The tasks were changed to align with expected learning in autumn 2004 and the new form of the laboratory manual was introduced. There was an apparently large increase in workload in this session (confirmed in the student evaluation general comments) that was directly attributable to completion of the learning tasks each week. Addressing the time demands of task completion presented a further issue for many students. Although this time demand was consistently decreased for subsequent sessions, students have persisted in believing the workload for this subject to be excessive. This has been evident not only in the completion rates for tasks as observed by the tutors, but also in the changes to the time recommendations for tasks listed in the laboratory manual. They have reflected the need for students to work beyond class time in order to complete them. In the light of continuing comment on the workload, it has appeared that the students regarded this expectation to be too demanding, believing that they should need to work only on assessment outside of allocated class hours.

6.6.3 The impact of the *laboratory manual* on student learning (see Appendix 6.3)

The *laboratory manual* contained the learning tasks that collectively modelled expected learning for assessment. There have been many changes to the *laboratory manual* across the implementations. With the introduction of the *manual* in its workbook form in autumn 2004, students recorded the task responses in it and task completion was encouraged by permitting students to take the completed manual into the final exam with them. Students perceived the benefits in committing to the work and despite the extra effort it demanded, students appeared to tutors to work far harder than they had in the past.

Books were checked weekly for completion, and solutions to tasks were released at the close of each week. Students consistently valued the worth of the *manual* (89-95%). However their belief in its importance led to an unravelling of its worth in terms of its promotion of student learning! It was anticipated that students would use the solutions to check their own work, but over the period from autumn 2004 and beyond, students short cut the task completion and secured copies of solutions from previous sessions. Its teaching/learning value diminished as 'cut and paste' solutions, rather than effort, propelled the task completion. Laboratory class attendance also deteriorated across this period.

6.6.4 The impact of the *laboratory tasks* on student learning (see Appendix 6.4)

Student perception of the importance of the *laboratory tasks* has increased in ranking since 2002 and in particular since 2004 (when the new *manual* was introduced). There may have been several reasons for this:

- defined objectives focused students on expected learning and these objectives *aligned* with corresponding assessment;
- *organisers* for *statistical thinking* aligned the tasks with the lecture material and formed the foundations of *scaffolded* thinking;
- tasks provided the *scaffolding* for student thinking that facilitated the transfer of the thinking to new contexts and modelling the thinking to be demonstrated in assessment;
- the *scaffolding* underpinned the criteria for marking of the corresponding assessment;
- the provision of white space for task completion and inclusion of a variety of tasks that anticipated learning, required reading and interpretation of output, and sought production and interpretation of output; and
- the inclusion of learning strategy tasks promoted synthesis and summary of what students perceived as disparate and complex concepts and procedures.

Again, even though these *mechanisms of alignment* and support for learning were present, only the most committed students recognised the structure:

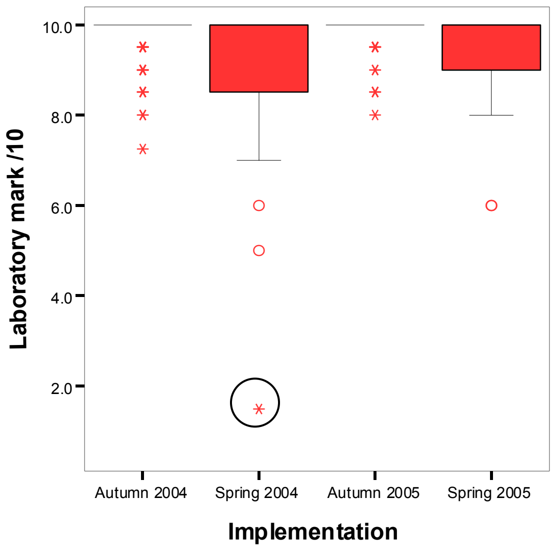
> Question: Do you have any comments about your perceptions of your learning in this unit? Please feel free to draw upon previous experiences for comparisons.
> Response: … The learning and assessment structure of Stat131 meant I always knew what was expected of me, and learnt the subject material over the 13 week period, and was obviously much more prepared for the exam. It also meant that I undertood the subject material, and as such, have retained a fair amount of what was taught … (Targeted STAT131 student's response to an online survey, autumn 2004).

Most students commented on the time required to complete the tasks, and this demand was progressively reduced in response to the level of concern:

> LESS LAB WORK!!!!!!!! Too much work to complete during the small amount of tutorial time!!! REVISE AMOUNT OF LAB WORK PER WEEK!!!!!! (A 'typical' student suggestion for improving the subject, autumn 2004 *student survey* data)
>
> Too much stuff crammed into 2 hour labs, lessen the amount of work !!!!!! (A 'typical' student suggestion for improving the subject, spring 2004 *student survey* data)
>
> some of the lab tasks take alot of time to complete, and there isnt too many instructions on how to use spss. so alot of time is spent using trial and error and then not sure that certian graphs have been made correctly or not. Assignments too long, too many parts to them for 10% assignments. (A 'typical' student suggestion for improving the subject, autumn 2005 *student survey* data)
>
> The Labs take up a lot of time , maybe reduce the amount of time required, the workload is heavy and when combined with other heavy workload subjects it is hard to do well in both subjects (A 'typical' student suggestion for improving the subject, spring 2005 *student survey* data)

The resultant distillation process has produced a more refined structure, with structured tasks that exemplify the desired thinking.

Completion of the laboratory tasks (reflected in the laboratory assessment mark) appears indicative not only of student perception of importance but also reflective of learning as evidenced in the final exam marks. Figure 6.4 gives the distribution of laboratory marks for all students 'passing' the final exam (exam marks greater than 50%). The median laboratory mark for all sessions for passing students was 10 (out of a maximum of 10 marks) and the means always exceeded 9. Across all sessions only one student with a failing laboratory mark passed the final exam (see Figure 6.4), and because of a debilitating medical condition, this student had secured an exemption from the compulsory completion of the learning tasks!



*Across all sessions, one student passed the final exam with a failing laboratory mark.

Source: Student assessment data

Figure 6.4: Distributions of marks for completed laboratory tasks for students achieving greater than 50% in the final exam

6.6.5 The impact of the *solutions* on student learning (see Appendix 6.5)

Student appreciation of the value of the *solutions* to *laboratory tasks* and *assessment* increased progressively across all of the surveyed sessions. However, this may be one instance where students do not necessarily recognise that they can have too much of a 'good thing'!

The teacher has long believed that only infrequent and judicious reference to worked solutions should be recommended in a mathematical learning environment. Such a position has resulted from her observations of successful learners and her own personal experience. Confident students only refer to 'answers' as a last resort, preferring to examine *what they know* and determine *what they still do not know*. This approach to learning appears to enhance understanding and promote a more robust grasp of concepts. Many weaker students on the other hand, attempt to replace the process of assimilation of new thinking with 'Oh yeah!' realisations founded on another's pre-structured response. However, the solutions rarely appear to bridge the gap between the student's inadequate grasp of the concepts and their acquisition of those concepts, thereby giving rise to the potential for the presentation of all problems as an infinite range of disparate questions each requiring equally disparate answers, an impossible learning task to master! Equally, offering 'answers only' to weaker students can promote a 'manufactured' logic founded on poor or absent understanding. Again, this rarely rectifies a learning deficit.

In the context of this subject, however, students attended two hour laboratory classes conducted as workshops where their efforts were facilitated by tutors. It was not initially anticipated (by the subject designer, the researcher or the teacher) that they would 'side step' the learning designed to take place in the laboratory classes. However, over this same period, laboratory attendance decreased, teachers reported that student effort in classes diminished and ultimately performance in the final exam did appear to suffer (see Table 6.5).

Table 6.5: Failure rates (%) across all implementations

Autumn 2003	Autumn 2004	Spring 2004	Autumn 2005	Spring 2005
10.7	9.4	15.0	20.4	23.1

Source: Student assessment data - grades.

A fresh approach to remedy this situation was instituted during 2006-2007, requiring students to sit mastery quizzes founded on the lab tasks. The following reflective comment was reported by the subject designer:

> … students increasingly read solutions rather trying to work the exercises, so in 2006 with the failure rate increasing, a system of laboratory tests based on the laboratory manual/work were introduced; this corresponded to a doubling of the high distinction and distinction grades, with students moving from primarily pass conceded and pass grades with the failure rate again starting to fall. (A general teaching evaluation commentary by the subject designer, 2006)

Since the completion of data collection for this study, the teacher has observed that the students have remained resistant to personal effort in completion for the first four weeks. However they eventually discover that *solutions* do not replace *active task completion* as their quiz results fail to meet a required 7/10. It is this realisation that has appeared to provide the necessary motivation to complete the tasks. In 2008, this approach has been extended to offer students a second attempt at demonstrating mastery, with unsuccessful students targeted for learning support.

6.6.6 The impact of the *assignments* on student learning (see Appendix 6.6)

Students generally worked on all three assignments with the same partner. Collaboration was enabled through *parallel* questions requiring partners to respond to similar questions using different data, and *complementary* questions requiring information to be extracted from a partner's responses. For example, one task might require a student to work with a theoretical approach to a binomial model while the partner might determine if data fitted that binomial model. In a later question, the roles might be reversed for a different probability model. For example, the student working with the binomial model might fit a Poisson model to the data using their partner's discussion of the theoretical implications of the Poisson model. Students responded variously to the demands of the teamwork required in the assignments. Many students failed to appreciate the team approach (see Table 6.6). Nevertheless the proportions perceiving a positive experience increased substantially from the Observation Phase in 2003.

Table 6.6: Percentage of students reporting that *teamwork worked well* across all implementations

Autumn 2003	Autumn 2004	Spring 2004	Autumn 2005	Spring 2005
8	12	49	53	38

Source: Student survey data.

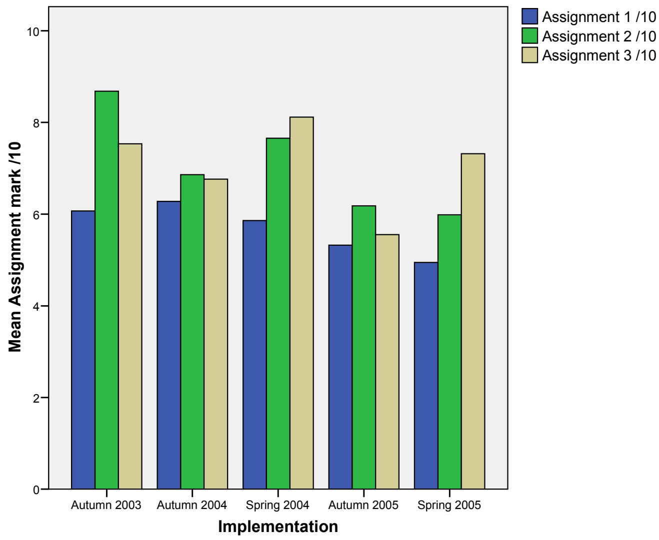
The students have demonstrated their belief in the value of assignments to their learning across all surveyed sessions, although their ranking of the importance did fall in 2005. However submission rates diminished across the sessions, particularly in 2005 (see Table 6.7). Perhaps this resulted from the students' lack of personal completion of the laboratory tasks and their consequent inability to access the *modelling* provided by those tasks.

Table 6.7: Submission rates (%) for assignments across all implementations

Task	Autumn 2003 (N=159)	Autumn 2004 (N=205)	Spring 2004 (N=172)	Autumn 2005 (N=179)	Spring 2005 (N=160)
Assignment 1	98.7	97.6	97.6	88.8	85.0
Assignment 2	95.0	89.5	89.5	82.5	82.5
Assignment 3	91.2	94.1	83.7	76.0	77.5

Source: Student assessment data

Mean assignment marks (see Figure 6.5 and Appendix 6.21) for the first two assignments decreased generally across all sessions from autumn 2004. The exception to this was spring 2004 which was modelled on the autumn session assignment for which solutions were widely available. Spring 2005 assignments were also modelled on autumn 2004. Assignment 3 marks show a similar pattern, except there is a distinctly higher mean mark in spring 2005.



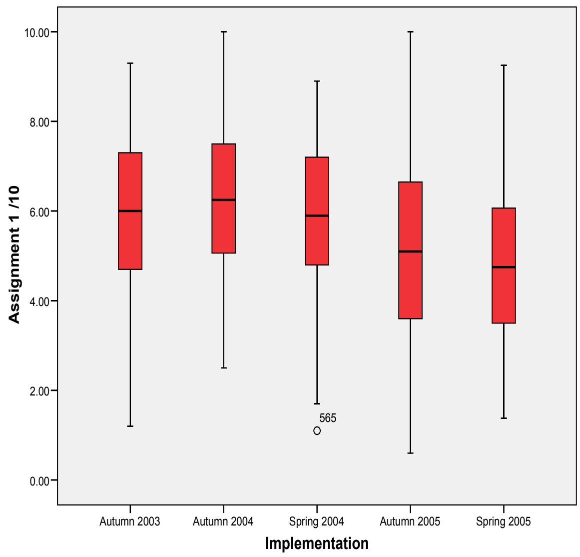
Source: Student assessment data

Figure 6.5: Comparison of mean marks for the three assignments across all implementations

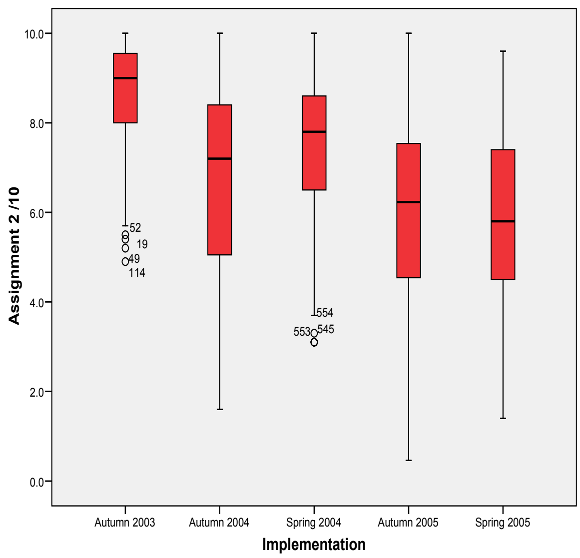
Following implementation of the teaching/learning framework, the first two assignments also became conceptually more difficult, addressing greater proportions of higher order cognitive processing. This increasing cognitive demand was also reflected in the corresponding questions in the final exams (Morris et al., 2007b). What is also evident in Figure 6.5 is an increase in the mean marks from Assignment 1 to Assignment 2 (within each implementation), (although again not consistently extending to Assignment 3) that is in keeping with student comments on their use of the *scaffolding* provided by the marking criteria after Assignment 1.

A comparison of the distributions of assignment marks however is even more confusing (see Figure 6.6). The distributions of the assignment marks from the observation phase are distinctly different from the other sessions for most of the assignments. The higher marks, smaller spreads and negative skew might correlate with a marking system focusing on recognition of lower order knowledge and skills. Among the sessions implementing the teaching/learning strategies, the means of the spring 2004 Assignment 1 marks differed significantly from those for the 2005 sessions. Again the cognitive demand increased for all assignments in 2005, with the incorporation of the ‘writing for meaning’ tasks known as ‘meaningful paragraphs’. For Assignment 2 in spring 2004, not incorporating this approach and modelled on autumn 2004, mean marks are distinctly higher. This pattern was also evident in the similarly modelled final exam.

In 2005, Assignment 3 incorporated an explanatory report on the student's data analysis. Hence these assignments required demonstration of complex statistical thinking and contextualised reporting on that thinking. Students find such writing tasks challenging and there were no exemplars available at that time. It should also be noted that approximately 20% of the enrolled students were international students whose first language was other than English and that writing tasks requiring contextualised explanations of complex discipline concepts may be especially difficult for many of these students.



The median (and mean) mark rises very slightly in 2004 but both median and mean marks then decrease across all sessions



The median and mean of autumn 2003 are significantly higher than all other sessions and the spread distinctly smaller.

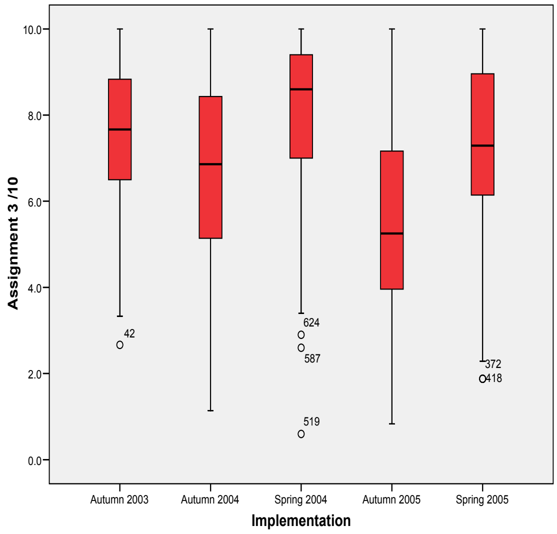
Among the implementation phases, there is evidence that all sessions are significantly different from each other except the two sessions in 2005.

There are some outliers at the lower extremes of the distributions of the implementation phases.

All sessions exhibit the negative skew that might be expected in a *mastery* task.

Source: Student assessment data

Figure 6. 6: Comparison of the distribution of assignment marks across all implementations



The median and mean of the 2003 observation phase are significantly different to the autumn sessions but not the spring sessions.

Among the implementation phases, spring 2004 and autumn 2005 are significantly different from all other sessions.

Outliers are evident at the lower extremes of most session mark distributions.

Source: Student assessment data

Figure 6. 6 (continued): Comparison of the distribution of assignment marks across all implementations

An example of an assignment has been included in Appendix 6.30.

6.6.7 The impact of the *lecture notes* on student learning (see Appendix 6.7)

The teacher/researcher observed that *lecture notes* appeared to form the basis for note taking during lectures, and to supply the discipline knowledge and modelling of procedures that were expected in completion of the *laboratory tasks*. In 2002, prior to restructure of this subject, 74% of students reported their value for *online lecture notes*, a rating of 7 (see Figure 6.3) among the surveyed facets of the subject's presentation. Although the proportions of students reporting the value of the *lecture notes* to their learning increased across the implementation phases, this increase was not accompanied by an equivalent change in the ranking. Students were also surveyed on their perception of the importance of the *online lecture notes* to their learning, but despite the two similar questions, ratings were not significantly different. One exception to this pattern was in autumn 2004 (see Appendix 6.17). However, the notes were also contained in the *laboratory manual* and the majority of students (92%) identified the *laboratory manual* to be important to their learning. This rated it as 2nd in terms of the proportion of students perceiving it important to their learning.

6.6.8 The impact of the *marking guides* on student learning (see Appendix 6.8)

The prototype of the *marking criteria* was piloted in the parallel subject (STAT151/252) in spring 2003. This testing highlighted the need to not only *align* the *objectives* with the *criteria* but also with the tasks themselves if it were to provide effective *scaffolding* for a student's *statistical thinking*. Marking too had been difficult, as marking time increased when the marker frequently had to search for the appropriate *criteria* to 'check'. The model implemented in STAT131 in autumn 2004 attempted to rectify these issues.

In rating the proportions of students reporting a facet of the teaching/learning framework as important to their learning, student responses (65-76%) did not result in a high ranking for the *marking guides*. The *marking guides*, however, were essential elements in *alignment* of assessment objectives and checking of student responses. They also provided *scaffolding* for responses, and their worth drew comment from the targeted student quoted in paragraphs 6.6 and also in 6.6.4:

> Question: Did you notice the 'marking guides' included in each assignment?
> Response: Not in the first assignment, but otherwise, yes.
> Question: Why do you think they were placed there?
> Response: So that students knew what a marker was looking for in a response.
> Question: Did you refer to them?
> Response: Not in the first assignment, which ended in a low mark. Further assignments were attempted keeping strictly to the marking guidelines, and ended in much higher marks. (Targeted STAT131 student's response to an online survey, autumn 2004)

This type response (although elicited from a student targeted for his commitment) might be more representative of the student cohorts than anticipated, as it was also consistent with the increases in marks between first and second assignments in all sessions (see Figure 6.5 and Appendix 6.21).

Student feedback on achievement, and associated perceptions of fairness in assessment provided further affirmation of the value of the criteria. Nevertheless, students did not rank the *marking guides* highly in their responses to importance to their learning.

Teachers commented initially on how labour intensive the construction of the criteria had been. However once in place amendment across sessions was facile, *alignment* was facilitated, and accountability strengthened. The markers were more positive in their responses to the *marking criteria*:

> 3. What do you believe to be the most successful facets of the subject's assessment regimen?
> Response: …Structure of marking scheme … (A marker of assignments in STAT131 in response to an online survey, spring 2005)

However this had not always been the case! In spring 2003, the marking guides for equivalent assignments were piloted in the 'parallel' subject, STAT151. The same markers were used for the tasks, however the 'structure' of the criteria aligned poorly with the questions and marking was both complex and time consuming. Relying on the markers' feedback (genuinely warranted grumbling), the researcher/teacher restructured the criteria and following its subsequent implementation in STAT131, marking proved very efficient. Student marked assignments were returned within seven days and both markers and students reported satisfaction with the process.

In autumn 2005, the *marking guides* accompanied the *assignments* as a *checklist*. In early assignments, they appeared before the tasks, while at the end of the session they were placed at the end of the tasks. Students invariably *checked* the lists, but when their assessment responses were marked, it was apparent that students had done this without reference to their answers and they had not used them to facilitate correct responses. The ranked importance of the *checklists* fell in 2005. Because of the non-representative and small sample in spring 2005, no conclusions have been drawn.

6.6.9 The impact of the *midterm* on student learning (see Appendix 6.9)

In 2002 and 2003, more than 80% of students reported valuing the *midterm* test as important to their learning. In these sessions, students constructed summary notes that they took to the exam. However, the proportions of students valuing the *midterm* test diminished from the beginning of autumn 2004, and even more sharply in 2005. The fall from favour was mirrored in its drop in ranked importance to learning from a ranking of 2 in 2003 to 13 in spring 2005. This might be surprising as the midterm marks were moderately (r=0.6) but significantly (p=0.01) correlated with the final exam marks. However the midterm was also similar in format, and hence requiring many similar skills. The midterm has since been removed from the assessment regimen, and has effectively been replaced by the regular laboratory quizzes based on the laboratory tasks.

6.6.10 The impact of the *online lecture notes* on student learning (see Appendix 6.10)

Students regularly accessed the *online lecture notes* while they were working on the *laboratory tasks*. However, the notes were also included in their *laboratory manuals* and hence all students had access to hard copies. During 2003, and autumn 2004, student focus moved away from the importance of the online lecture notes, with associated falls in rankings (see Appendix 6.17), toward facets that engaged them more actively in learning: *assignments*, *laboratory classes*, and *tasks* (including their associated *solutions*).

6.6.11 The impact of *teamwork* on student learning (see Appendix 6.11)

Teamwork was not ranked highly by students in the sessions surveyed. Nevertheless it is an espoused graduate attribute for the University of Wollongong and is highly prized by employers and the general community (West, 1998). Students regarded the marking of the team assignments fair, believing that

> i'd say it is extremely difficult to receive marks without earning them. Assessment system was fair. I think that the assignment parts a and part b were basically equivalent in difficulty. (A student survey response, autumn 2004)
>
> The marking system is fair, students who put in the work get the marks and those who don't put in the work don't achieve many high marks at all. (A student survey response, autumn 2004)
>
> … because it isn't necessary that two people have to complete the assignments together. It is fair because each person is equally responsible

> for getting the assignment complete and therefore should be marked based on their contribution to the assignment. (A student survey response, autumn 2004)

While a few students still appeared to believe that despite teacher effort, in some teams the workload was uneven, general comments from 2004-2005 did not elicit a single negative comment about teamwork. Nevertheless, even though the rates generally increased across the sessions, fewer than 53% of students reported a positive attitude to teamwork.

6.6.12 The impact of the *tutor* on student learning (see Appendix 6.12)

The proportions of students perceiving tutors as important to their learning were similar for most sessions (70-80%) with the distinct exception of except spring 2004 (21%). The rating of these proportions remained stable. While some general comments in the spring 2004 survey reflected negative attitudes towards their tutor, little commonality existed between them. However, the comments widely reflected dissatisfaction with the 'timing' of release of *assignments*, *solutions*, and *lecture notes*, and this may have been deflected to the tutor who was the most 'available' teaching contact. These timing issues had not been anticipated in the 'handover' from the subject designer to the spring coordinator, highlighting the necessity for documenting such timelines in order to maintain the *alignment* of the learning framework.

Also at this stage, the tutors were not generally issued with worked *solutions* for the *laboratory tasks* and this may have made quick resolution of exercise task difficulties difficult. Interestingly in the 'parallel' subject (STAT151/252), most students commented that the *laboratory classes* were extremely important in their learning and affirmed value for both the *tutors* and the *classes*.

The following comment, although indicative of a negative attitude to the tutor, recognised the importance of the laboratory exercises and its connection to the lectures:

> Practical explanation was lacking. I find that solving the problems yourself is the best way to learn how to do them. The lab is the most important part of understanding these concepts and the tutor needs to listen to questions from an individual then get the entire class' attention and show how to work it out, not just show the student who asked. Lectures may explain the theory

> but until you are challenged by a problem to solve, the theory is meaningless. (Student survey comment, spring 2004)

Student comments on the necessity of engagement between the *tutor* and students, and the *laboratory classes* and the *lectures* have promoted increased effort in tutor support; more structured facilitation of student learning in this arena; and improved *alignment* between the lecture material and the laboratory exercises.

6.6.13 The impact of the *learning strategies* on student learning (see Appendix 6.13)

Although the *learning strategies* tasks were not incorporated into the *laboratory exercises* until spring 2004, they were discussed in the *laboratory manual* introduction and modelled by the lecturer and the teachers in the *laboratory classes*. Such models also provided frameworks summarising discipline knowledge in the *solutions* to the laboratory tasks from autumn 2004. From the teacher's observations, few students were able to devise such strategies or even to complete a framework without teacher support. However, once completed, they enabled understanding and provided mechanisms for transfer of the knowledge and skills. When commenting on '…ways to improve this subject…' in the student evaluation survey, the following student makes reference to the learning strategies:

> I think more 'tying-together' of the subject material would help, like what we do in the mind-mapping / framing tasks - but done for us both before entering a topic and as a recapitulation when leaving a topic. Also, maybe presenting the forumula's in a few different ways so we are aware of how they might be shown in different books, and tying together the common parts of formula and explaining them as concepts not algebra could be helpful. (Student survey comment, spring 2004)

However this appeared to be a solitary voice. There was a steady improvement in the ratings of the proportions reporting value of the *learning strategies* to learning, and since 2005 their increasing prominence in the learning tasks has also increased student recognition of this value.

6.6.14 The impact of the specified *objectives* on student learning (see Appendix 6.14)

For the teacher, the *objectives* have proved to be the primary mechanism of *alignment*. The other participating teachers have also commented on their promotion of focus for teaching. Students have consistently made reference to the aligned structure of the subject

> It is the best ran and most well organised subject that I have ever taken. Maybe some summary sheet of rules & formulas - although i realise that there is much to learn from constructing my own. (Student survey comment, autumn 2004)

Targeted student survey responses (see paragraph 6.6) have also acknowledged the potential value of the *objectives*, but this has not translated to their personal perception of the value of the objectives to their learning. Perhaps in the end, it is more important that the teachers understand their value, for they are the agents for implementation of the aligned pedagogy. It is not necessarily important that students be able to identify the mechanism, but more important that they perceive the *alignment of intent and practice* and the connective structure that relates all aspects of the subject's presentation. This has been amply evident in their responses to fairness and in their general comments.

6.6.15 The impact of the online *forum* on student learning (see Appendix 6.15)

The subject designer encouraged students to discuss their problems openly in the online *forum* and indeed there was intense communication, particularly before assignments. The designer also regularly interacted with students in this way; however survey responses do not indicate that many students value this form of communication as important for their learning, which was ranked at or near the bottom of the facets surveyed each session.

6.6.16 The impact of the *text book* on student learning (see Appendix 6.16)

Few students regarded the *text book* as important to their learning. There was some value indicated by students in spring 2004. The spring session coordinator did make reference to it during lectures and students responded accordingly. However, since there was no mechanism developed to *align* the *text book* to the other facets of the subject's presentation, few students managed to incorporate it into their study regimen independently. Most students appeared to regard the *lecture notes* as defining the content (a role frequently ascribed to a *text book*). The subject designer no longer defines a *text book* for this subject.

6.6.17 Students' general comments.

Throughout the previous discussion, continual reference has been made to student comments. The teacher has previously declared her view that students *know what they know* and *know what they don't know*. Their considered reflective comments on their learning in this subject have not only affirmed the teacher's view but that of the researcher. Their comments have propelled improvements in *alignment* and identified many other strategies for improving their learning that have been adopted for subsequent implementations:

> I think the biggest improvement should be the timeliness of putting stuff up on the WebCT. I found it useful to have the lecture notes in front of me whilst attending the lectures, however this was not always possible, because they weren't always posted to the WebCT before a lecture and indeed sometimes not until days after the lecture …. Lectures may explain the theory but until you are challenged by a problem to solve, the theory is meaningless. Even when (the subject designer) taught this subject in Autumn semester and went to great lengths to describe real life applications, the "understanding" did not come until I actually had to solve the problem myself. It was at the point where I was asking the questions about how to solve the problem that I absorbed and retained an understanding of what was being taught. … The assignments were also made more difficult by not having the relevant Lab Solutions posted to the WebCT on time.(Student survey comment, spring 2004)

> formula's explained better and better review in our lab times as review was rarely completeed in my lab (Student survey comment, autumn 2004)

Student comments widely reflected dissatisfaction with their workload and the self-directed nature of the exercises, believing that these limited their capacity to understand the concepts and procedures. In the spring sessions, students were given access to support tutorials, and although the numbers attending were not large, the students expressed their value for these tutorials that were in essence a return to the comfort zone of high school classrooms. This attitude had been even more evident in the parallel subject, where alignment was an issue and attendance at these support tutorials was more substantial.

6.7 Tracking the evidence: overall learning

The researcher examined student learning:

1. from their own perspectives in the student survey responses to topic learning, graduate attribute achievement and overall learning;
2. from the teacher's perspective through exploration of the assessment marks; and
3. through classification of the assessment questions (and answers) using the revised taxonomy of Bloom (Anderson et al., 2001).

The ranked proportions of student perceptions of topic learning were also correlated with their topic achievement in the final examination in order to test the contention that students *know what they know* and *know what they don't know*.

6.7.1 The students' perspective: discipline learning (see Appendix 6.19)

For spring 2004 to spring 2005, the student survey sought perceptions of confidence in 'topic' learning:

1. use of SPSS;
2. basic data exploration;
3. exploring relationships between variables;
4. resolving problems related to discrete distributions;
5. resolving problems related to continuous distributions;

6. checking to see if data fit a model;
7. setting up and interpreting confidence intervals;
8. setting up and interpreting hypothesis tests;
9. using a Markov chain to calculate probabilities; and
10. identifying the appropriate formulae to solve problems.

Summaries of the proportions of responses claiming confidence in their topic learning are given in Appendix 6.19. In spring 2004 the researcher deconstructed the students' final exam marks into the surveyed topics. It should be noted that only seven topics were tested in a way that enabled isolation of the marks. Use of SPSS and identifying formulae could not be distinguished from more theoretical discipline targets and no questions involved continuous distributions (neither Normal nor Exponential). This possibly impacted upon student performance in comparison to other sessions, as students find questions dealing with these distributions challenging. The correlation between the mean marks achieved by students in these topics in the final exam with the proportions reporting confidence in topic learning(see Appendix 6.19) was significant ($p<0.02$) and strong ($r=0.9$) (Morris et al., 2004b).

There is no linkage possible between individual student assessment results and responses by the same students to the survey questions (except for a small and potentially biased sample of students who gave permissions for such a link). Therefore, any conclusions based on associations relate to the student body as a whole rather than to individual students and hence fall short of supporting evidence for the contention that students *know what they know* and *know what they don't know*. However, these results allow speculation that any association found will also apply to individuals, although the strength of association may be less.

6.7.2 The students' perspective: graduate attributes (see Appendices 6.20 and 6.26)

The University of Wollongong has recently revised its specification of the graduate attributes, now referred to as the graduate qualities. They describe qualities the university promotes in students through its teaching and learning:

- Informed;
- Independent learners;
- Problem solvers;

- Effective communicators; and
- Responsible (University of Wollongong, 2008).

It has been a part of the teaching brief to develop these qualities in the context of statistical thinking in the subject STAT131.

In both sessions in 2005, the student survey asked students about their progression towards achieving the attributes (see Appendix 6.25, questions 40 – 46 for detail of the questions). Appendix 6.20 details the proportions of students claiming progression towards achievement of the survey attributes. Appendix 6.24 relates the surveyed attributes to the above graduate qualities.

There is too much dissimilarity, both in the responses and profiles of the students responding (see discussion of potential bias in spring 2005 survey responses in paragraph 6.5.3) to expect reliability in comparison between the two sessions. Common to both sets of responses, however, is the high ranking in terms of percentages reporting progress in working effectively and responsibly in teams. Teamwork was a quality that was directly targeted in the collaborative learning strategies underpinning classroom learning and also the formative assessment (in particular the assignments). This might appear to be at odds with the low ranking of percentages reporting the ‘team approach’ as important to their learning (see paragraph 6.6.11). The percentages however are not too dissimilar, and they may also reflect the increased demand upon student time in working as part of a team.

Because of potential bias in the small sample in spring 2005, it is more relevant to the current discussion to refer to the autumn data alone. In autumn 2005, the greatest proportion of students reported progress in the use of technology. This was not surprising as few students claimed to have used a statistical software package for statistical analysis.

Associated with the belief in their ability to appropriately use technology was their perception of progress (55% of students, see Appendix 24) in their ‘…ability to use technology to analyse, organise and present data as information.’ This may in part be attributable to their increased expertise in using SPSS to perform basic exploration. However, it also incorporates more complex learning that requires selection of appropriate discipline techniques/procedures, in keeping with the assumptions/criteria underlying those techniques/procedures.

Allied to this reported progress was the students' belief in their ability to '… make choices in the analysis of data and logically justify these choices' (53% of students). It should also be remembered however that the teaching/learning framework for this subject focused on the development of evidence based decision making and which in turn required development of analytical skills.

6.7.3 The students' perspective: overall learning

From autumn 2004 to spring 2005, students were surveyed about their perception of exam preparedness as a consequence of completing the subject requirements (see Appendix 6.25, question 24), their belief in their overall statistical learning (see Appendix 6.25, question 25) and their perceptions of subject relevance (see Appendix 6.25, question 26). The results have been tabulated in Tables 6.8 -6.10. Again learning confidence is strongest in spring 2004, with the greatest proportion of students believing themselves prepared for the final exam (56% reporting preparedness or only a need for basic revision) (see Table 6.8) and the strongest belief in their overall statistical learning (64% reporting at least moderate success in learning) (see Table 6.9). Most students do not report improving their perception of the subject's relevance to their lives (Table 6.10).

Table 6.8: Percentage of students reporting belief in exam preparedness after completing all set tasks across all surveyed sessions

Response	Implementation			
	Autumn 2004	Spring 2004	Autumn 2005	Spring 2005
Too much to learn	28	19	26	33
Still need revision	30	25	18	24
Need basic revision	31	27	27	14
Well prepared	9	29	27	29
No response	2		2	

Source: Student survey data

Table 6.9 : Percentage of students reporting belief in *statistical learning* across all surveyed sessions

Response	Implementation			
	Autumn 2004	Spring 2004	Autumn 2005	Spring 2005
Too difficult	18	3	10	5
Tried unsuccessful	28	14	32	24
Tried limited success	37	19	16	24
Tried moderate success	12	54	30	42
Learned a great deal	3	10	10	5
No response	2		2	

Source: Student survey data

Table 6.10: Percentage of students reporting perception of subject relevance across all surveyed sessions

	Implementation			
	Autumn 2004	Spring 2004	Autumn 2005	Autumn 2005
Less relevant	21	17	29	28
Same relevance	47	56	43	43
More relevant	30	27	23	29
No response	2		2	

Source: Student survey data

6.7.4 The teacher's perspective: the marks

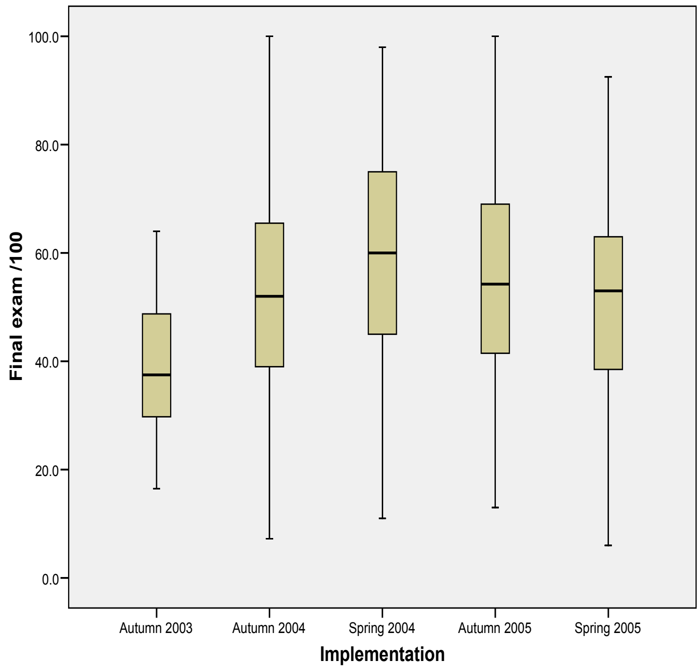
The researcher has operated from the position that the final exam should target the desired learning as espoused in the subject's learning objectives. The teaching and learning frameworks have been designed around these learning outcomes. Some might criticise that the teaching has been tailored to the test. This might be criticism indeed if the test

> … elicit (s) lower level cognitive activities than the objectives nominate …
> (Biggs, 1999, p. 68).

However the assessment has been aligned with the objectives in terms not only of required discipline knowledge of terminology and procedures, but also in terms of cognitive demand.

> So, if students focus on the assessment, they will be learning what the objectives say they should be learning. (Biggs, 1999, p. 68)

Hence in an aligned teaching learning framework, in determining the extent to which learning has been taking place, it is important to investigate the final exam. This is done from two perspectives: student achievement as revealed by their exam results (in this paragraph) and from the exams and the cognitive demand of the questions revealed in the application of the revised taxonomy of Bloom (Anderson and Krathwohl, 2001) (see paragraph 6.7.5).



Source: Student assessment data (zero final exam marks removed as these indicate non attendance rather than nil achievement).

Note: The variation in autumn 2003 (pre-implementation) is distinctly different to the other sessions and hence post hoc tests of pair wise comparisons should account for this (see Appendix 6.22)

Figure 6.7: Boxplot of final exam marks across all sessions

Using ANOVA to test for differences between the means of all sessions, has indicated significant differences between the mean final exam marks ($p<0.001$). Removing the Observation Phase (autumn 2003) from consideration, a comparison of the means from autumn 2004 to spring 2005 highlighted significant differences between the means ($p<0.001$) and post hoc comparisons (see Appendix 6.22) identified spring 2004 as significantly different from all other sessions ($p<0.04$) with no significant differences observed between pairs of the other implementation sessions.

The reasons behind the significantly different spring 2004 data have been difficult to isolate. Critical review has identified a number of relevant issues:

1. There were logistical problems not identified prior to the hand over of the subject from the subject designer to the spring coordinator. These predominantly revolved around scheduling of assignments in alignment with associated laboratory learning (source: annotated journal);
2. One important issue was the late release of relevant laboratory task solutions beyond the due dates of assignments. Students were unable to address learning deficits from the laboratory tasks before they were required to submit solutions to the related tasks (source: annotated journal and student survey comments);
3. The distribution of laboratory marks for this session was different to all other sessions (see paragraph 6.6.4);
4. Few students recognised the value of the tutor to their learning (21%) (see Appendix 6.12);
5. High ranking of importance of assignments and laboratory manual by the students (see Appendix 6.17);
6. Greater proportions of students reporting ‘preparedness for the exam’ (see Table 6.8) and greater perception of overall statistical learning (see Table 6.9);
7. Although there appears to be little difference in mean assessment marks for early assessment, later tasks have apparently higher means (see Appendix 6.21 and Figures 6.5 and 6.6);
8. The coordinator modelled assessment on the autumn assessment tasks, changing only the data sets and hence students had access to previous exam questions in their preparation (source: annotated journal);
9. The final examination included no questions on continuous distributions (in particular the Exponential distribution) and students traditionally find such questions challenging; and
10. Student perceptions of topic learning were higher (see Appendix 6.19).

The first four points might be expected to predicate a ‘downturn’ in marks. However as Figure 6.7 shows, and the results of testing for differences between the means has revealed, the mean mark for the spring 2004 exam was significantly higher. Points 5 to 10 above do lend credence to this result.

6.7.5 From the researcher/teacher’s perspective: the exams (see Appendices 6.26-6.29, 6.31a-c)

If the final exams are designed to detect achievement of the desired learning, then shifts in student learning should be evident in their marks for these examinations. The researcher needed to determine if the shifts in the student marks were merely manifestations of shifts in the level of knowledge and skills demanded by the examination questions, or whether the teaching/learning framework had promoted more complex cognition. In order to do this she has

- coded the marks allocated to questions in the exams using the revised taxonomy of Bloom (Anderson and Krathwohl, 2001);
- split the knowledge types and cognitive processing skills into ‘higher order’ and ‘lower order’. Her allocation to these categories has been grounded in her experience in the classroom, affirmed through consultation with peers and confirmed subsequently in the literature (see Chapter 3, paragraph 3.4.2); and
- compared percentage representation of the two categories across the implementations.

Only the autumn sessions were examined as these were designed by the subject designer, whereas the spring sessions were modelled closely on the autumn 2004 session (although the spring 2004 exam omitted some of the more challenging topics e.g. no questions on continuous distributions). In addition, only the questions related to *bivariate distributions/correlation* were selected for classification. The researcher selected this topic because

- she believed that the teaching/learning framework had impacted on student learning in this topic;
- she believed that assessment of this material had increased in cognitive demand in this topic; and

- the topic was examined in a single question and hence the marks targeted for the target were readily identified.

The results of the investigation have been tabulated in Appendices 6.26 – 6.29 and summarised in Table 6.11. The percentage of marks allocated to higher order knowledge type questions increased in the first implementation (autumn 2004) although it decreased in autumn 2005. Reference to the detail in the appendices however will highlight the large number of marks allocated in 2005 to *Procedural skills and techniques,* whereas in 2003 there was a greater percentage allocated to basic skills of *Remembering.* There were other differences in the exams that have proved difficult to capture in the classification. The questions in the autumn 2005 exam were more textual in nature than either of the other exams. With such a large percentage of international students many may have found deconstruction of the questions difficult and this may have impacted upon the results generally.

Table 6.11: Percentage representation of cognitive demand[1] of final exams across autumn 2003, autumn 2004 and autumn 2005[2] (regression question)

Target of marks	Autumn 2003		Autumn 2004		Autumn 2005	
	Higher order[3]	Lower order[4]	Higher order[3]	Lower order[4]	Higher order[3]	Lower order[4]
Knowledge	30	70	40	60	10	90
Cognitive skill	40	60	60	40	75	25
Total higher order targeted[5]	45		80		75	

[1] See Appendix 6.29 for split in level of knowledge and skill types as classified by the revised taxonomy (Anderson and Krathwohl, 2001)

[2]Compiled from Appendices 6.26 to 6.30

[3] A higher order component (either knowledge type or skill) targeted in the mark

[4] A lower order knowledge component (either knowledge type or skill) targeted in the mark

[5] A higher order component (either knowledge type or skill or both) to the question.

Although many statistical procedures are complex, the focus in this subject has been evidence based decision making and contextualised commentary. The complex procedural calculations have been undertaken using the software package SPSS. This however does not diminish the cognitive demand of producing and identifying relevant output and assembling evidence, and it is this *thinking* that has been targeted in the teaching!

There is strong evidence of a shift in the demand for higher order cognitive processing skills from the observation phase in autumn 2003 (see Table 6.12) and in the overall level of demand. This analysis would seem to indicate that the exams have become more challenging following the implementation of the teaching/learning framework, perhaps a reflection of increased learning expected by the designer and the teacher! When considered in conjunction with the marks for the final exam a shift in learning does indeed appear to have taken place. One less impressive side effect, however, might have been an increase in the spread of the marks - a challenge for future implementations!

6.8 Conclusions:

The aim of this study has been to design and evaluate a teaching/learning framework that promoted achievement of the subject objectives, in particular *statistical thinking*. The evaluation encompassed analysis of student attitudes to the teaching/learning framework and perceptions of their learning. Teachers and marking staff were canvassed for their comments on the framework and their observations on student learning within that framework. As summative images of student learning, student exam results were analysed and the questions deconstructed to determine their level of cognitive demand.

6.8.1 Achievements

There has been evidence of improvements in students' *statistical thinking*. This has firstly been demonstrated in the student surveys. They generally have responded positively to the *alignment* of the teaching and learning, although they have not necessarily demonstrated appreciation of the mechanisms of that *alignment*. This subject is a foundation course in statistics, and yet more than half of students surveyed believed that they had made progress in achieving graduate qualities (particularly those reflecting the capacity to build evidence-based argument). Examined in the light of their performances in summative assessment, this has spoken strongly of increased student confidence in their learning in this subject. This evidence has also supported the contention that students *know what they know* and *know what they do not know.* This awareness of their learning, the learning strategies built into the laboratory tasks and assessment, and the commitment evidenced in the hours spent out of class have also been indicative of increased *meta-cognition*.

Teaching and marking staff have commented on the focus that an objective driven approach has given to their teaching/marking. Aligned curriculum has enabled the promotion of more complex learning and this has been evidenced in the exam questions themselves and the student results in those exams. Higher order thinking has been addressed by the framework and achievement demonstrated in the final exam results, with the mean marks rising following the framework's implementation and increased cognitive demand in the examination questions.

Student survey responses and their comments indicate awareness of *fairness* has been generally perceived in assessment, and accountability enhanced with the strict *marking criteria*. Markers reported that marking was expedited by the *alignment* of the *marking criteria* with the tasks and the dichotomous approach to marking. Students reported positively on the feedback provided by the 'checked' *criteria*.

For the teacher, her evaluation has proved enlightening. This research has inspired in her a sense of awe at the teacher's instinctive ability to react to students and to evaluate their learning. The only limiting factor in her understanding the action in the classroom has been organising the wealth of evidence that lies at hand!

6.8.2 Deficiencies highlighted during the implementation

Evaluation has detected a number of logistical issues that proved important to student attitude and learning:

- Handing over to a new subject coordinator highlighted the need for strict scheduling of tasks and release of solutions if feedback is to be optimised;
- Students learned to short circuit learning by *downloading solutions* rather than completing laboratory tasks;
- Students preferred no *common* questions on assignments. Different questions increased their perceptions of *fairness;* and
- *Alignment* of the *marking criteria* with the tasks expedited the marking process.

In terms of extending learning and assessing that learning

- greater awareness could ensure representation of more diverse higher order knowledge types in the final exam;
- increasing the textual demand of questions creates issues related to its deconstruction for meaning for some students; and
- *alignment* of the *marking criteria* with both the *objectives* and the tasks provided more effective *scaffolding* for *statistical thinking*.

The researcher has found reliable application of the revised taxonomy difficult in this subject. She has repeated her classifications several times and conferred with peers in order to ensure appropriate classification of the exam questions, but the process has identified the need for a discipline based classification of statistical learning.

6.9 New paths to tread

As with all teaching environments, impacting variables change with each implementation and it is the role of the designer/teacher to respond to these ever changing demands. No teaching/learning framework sits comfortably in all situations. However the framework developed within this study has proved elastic to the needs and ambitions of its protagonists. There have been some areas however that have arisen during the study and the researcher has targeted them for further work.

6.9.1 *Peer review* and its impact upon *meta-cognition*

As a consequence of her work in the University of Western Sydney study (see Chapter Seven) the researcher has piloted attempts to utilise *peer review* in her promotion of *statistical thinking*. Again the *marking criteria* have provided the *scaffold*, but initial observation has indicated that the reflective process of a guided and evidence based evaluation of another student's thinking, might effectively promote *meta-cognition*. It has appeared to provide the overlayed cycle of *learning to learn* described by Watkins et al. (2002) and illustrated in Chapter 3 (figure 3.2).

As students in this subject submit at least two assignments and complete a presentation, there exists a possibility for marking a *peer review* in replacement of one of these tasks. Evaluation of implementation of such a practice could provide more substantive evidence of its effect on student learning.

6.9.2 Timing is of the essence!

Two issues revolving around timing have arisen in this study. The first is the importance of scheduling classroom tasks and solution release in order to provide effective feedback relevant to assessment. This was identified in the handover to the alternative subject coordinator. The second issue is the lifespan of some strategies engaged to guide student learning. Increasingly throughout this study, the students replaced *active engagement* in learning with *downloading of solutions.* This led (as might be expected) to a downturn in performance in assessment. This issue has been addressed by instituting a system of laboratory tests designed to check mastery of the learning developed in the weekly laboratory tasks. Students have proved resistant to the notion of mastery testing in the first four weeks, but eventually commit to working through the tasks actively. There has been an improvement in assignment marks and in student performance in the final exam. This approach to engagement is in need of more rigorous evaluation.

6.9.3 Classifying learning: working backwards

Much of the evidence for concluding that there have been improvements in student learning and more particularly, *statistical thinking* since implementing the teaching/learning framework has hinged on shifts in cognitive demand in the final exams and a simultaneous positive shift in mean marks in the final exams. This has required more effort than it might have had classification of the questions been more reliable. The generality of the revised taxonomy (Anderson and Krathwohl, 2001) has made it difficult to implement. Notwithstanding the work of Chance (2002), Garfield (2002), delMas (2002) and Gal and Garfield (1997), the classification of *statistical thinking* is in need of refinement.

Just as assessment has been the starting point for design of the teaching/learning framework for this study, so too could student responses to assessment provide us with a mechanism for classifying their thinking. Commonality of student assessment responses may yield an avenue for statistical analysis that would enable classification of the order of statistical learning. Text deconstruction of the questions could facilitate segregation of different orders of knowledge and skills peculiar to *statistical learning*.

Chapter Seven
Case Study
Making Sense of Nonsense - an excursion into critical thinking

"I dare say you are wondering why I don't put my arm around your waist," the Duchess said after a pause: "the reason is, that I'm doubtful about the temper of your flamingo. Shall I try the experiment?"

"He might bite," Alice cautiously replied, not feeling at all anxious to have the experiment tried.

"Very true," said the Duchess: "flamingos and mustard both bite. And the moral of that is – 'Birds of a feather flock together.' "

"Only mustard isn't a bird," Alice remarked.

"Right, as usual," said the Duchess: "what a clear way you have of putting things!"

"It's a mineral, I *think*," said Alice.

"Of course it is," said the duchess, who seemed ready to agree with everything that Alice said; 'there's a large mustard-mine near here. And the moral of that is – 'The more there is of mine, the less there is of yours.' "

"Oh, I know!" exclaimed Alice, who had not attended to this last remark. "It's a vegetable. It doesn't look like one, but it is."

"I quite agree with you," said the Duchess; "and the moral of that is- 'Be what you would seem to be'- or if you'd like it put more simply- 'Never imagine yourself not to be otherwise than what it might appear to others that what you were or might have been was not otherwise than what you had been would have appeared to them to be otherwise.' "

"I think that I should understand that better," Alice said very politely, "if I had it written down: but I can't quite follow it as you say it."

From The Mock-Turtle's Story in
Alice in Wonderland by Lewis Carroll, p.78

7.0 Background

7.0.1 Uncharted territories

The researcher had presented her work on the use of defined learning objectives to structure marking criteria that scaffolded expected learning to colleagues on many occasions. One instant response was the claim that the approach was more likely to 'sit comfortably in sequential and process driven disciplines such as mathematics and statistics'. The teacher had spent many years in classrooms trying to teach students to use discipline knowledge and skills to *think* mathematically/statistically. Her teaching intent was more closely targeted at promoting adaptive *thinking* to solve problems rather than procedural repetition of discipline processes. Her colleagues' comments provoked this case study as the researcher attempted to address student acquisition of critical thinking and evidence-based evaluation skills in a subject requiring deconstruction and critical evaluation of ideas grounded in text and not symbols. The researcher holds no expertise in the discipline associated with the subject; she was engaged to collaborate with the teaching staff because of her research and proficiency as a teacher.

7.0.2 Answering the critics

A common criticism of behaviourally defined objectives is that they set *limitations* on student learning. However, the objectives merely enable specification of a reference standard for achievement, and it is the learning tasks themselves which should afford scope for demonstration of learning. Although assessment may only *check* for achievement of specified objectives, the tasks themselves should entice students to always work to their highest level of ability. Specification of objectives does, however, facilitate alignment of teaching, learning and assessment (Biggs, 1999).

Even with clearly specified objectives, it can be presumptuous to assume that an assessment regimen will adequately capture the desired achievement of both knowledge and skills. This is particularly true when the skills addressed by assessment involve critical thinking. Evaluation of the theories and philosophies of others is highly desirable in most disciplines, but may present a challenge even for many academics! For students, many of whom have survived university by absorbing and re-using the ideas of others, critiquing an argument may prove extremely difficult. The question must then be asked: what type of learning framework would best lend itself to acquisition of such a prized skill? Devising an answer to this question has formed the focus of this case study.

7.0.3 A dilemma: addressing shallow learning in accounting students

Teaching staff from the Faculty of Business at the University of Western Sydney had been concerned at the level of achievement of their final year students in an undergraduate accounting degree. The subject was Accounting Philosophies and Theories and it introduced students to a variety of methodological approaches to analytical and critical thinking in the accounting discipline. Students were generally perceived by teaching staff as exhibiting *surface* knowledge of the subject content but many failed to demonstrate achievement of the *deeper learning* that was desired. They apparently found it difficult to critically appreciate the theoretical perspectives presented to them. Indeed many could not effectively structure and substantiate arguments in either oral or written form.

7.0.4 Setting aims for the study

The researcher was engaged to design a teaching/learning framework that would facilitate student achievement of the lofty learning objectives of the subject. Teachers wanted students to develop a *deeper learning* which would enrich their appreciation of the accounting discipline and endure through future experience. The researcher was engaged to develop classroom strategies which might facilitate achievement of those aims. As the researcher had no expertise in the discipline, she needed to engage in lengthy discussions with the teachers to effect the development of an appropriate teaching/learning framework. The researcher chose to use assessment as an initial focus for discussion of the desired learning outcomes and hence the formulation of the subject objectives (Biggs, 1999).

With this in mind, intensive discussion with the primary teaching staff centred on:

- articulation of the desired learning;
- determination of *behavioural hallmarks* of this learning;
- subject presentation to address the required knowledge and skills; and
- assessment targeting the desired learning.

7.0.5 A methodology

This case study encompasses a single implementation of a constructively aligned curriculum targeting student development of critical and evaluative thinking in an undergraduate subject for accounting students. The study has followed a grounded research approach, beginning with the definition of the desired learning through extensive discussions with the teachers. Using a mixed method approach, data included student survey responses, student and teacher comments, assessment results and the researcher's annotated journal. A review of the literature served to ground the pedagogical design in current theory and practice, and discussion and reflective practice informed revision prior to implementation. Evaluation referred to all data in determining if the implemented teaching/learning framework had enhanced student acquisition of the desired knowledge and skills.

7.1 Environmental detail

The University of Western Sydney is a large university with encompassing multiple campuses, most in the western and south western regions of Sydney, Australia.

7.1.1 The teachers

The teachers of this subject were committed to improving student learning, and although the teaching of this subject had been reviewed previously, there had been no attempt to align the entire teaching/learning framework. The teachers were regarded as experts in their discipline, both as researchers and practitioners, but had sought help in addressing their students' apparent inability to critically analyse and evaluate recent ideas and theories associated with their chosen discipline.

7.1.2 The students

The subject around which this case study was centred, was a compulsory subject (involving 223 students) in a degree program with a large student enrolment (6000 students across all years). It was delivered across several of the university's campuses.

Most students enrolled in the subject were in their final year of study in the three year degree program in accounting.

7.1.3 The subject delivery

Teaching for this subject involved delivery of large group lectures supported by smaller tutorial classes. Lectures afforded the most expeditious means of disseminating the bulk of the information when all campuses administered the same assessment.

The tutorial classes of about (30 students) were traditional in teaching design, revolving around *tutor facilitated* discussion of student responses to focus questions prepared by students in advance. Relevant articles were placed on the web and students were expected to read them and use the detail to support their submissions. Prepared responses were submitted for marking on a *done/not done* basis. Student centred discussion frequently degenerated into a *tutor led exposition* with students as *passive recipients*. A participation mark reflected student attendance and class preparation, a practice that has since been amended to more appropriately reflect participation and effort.

7.1.4 The assessment regimen

Students' final marks were an aggregate of marks awarded for them:

- an essay;
- a group presentation;
- tutorial participation and
- a final exam.

Although the subject designers were cognisant of their objectives for the assessment regimen, several dilemmas presented themselves in the practice of grading student responses:

- the students themselves appeared to be confused about what was expected of them;
- 'value judgements' were made as to the level of achievement within a range of marks, a practice with inherent problems of perceived *fairness*; and

- reporting the level of achievement to students offered little to help them recognise how to remedy their difficulties and hence improve their grade.

7.2 Defining the learning

It was the researcher's first task to facilitate the teachers' clear definition of the learning they wanted to see in their students as a result of student engagement in this subject. This learning not only included discipline knowledge and skills but the generic skills that enable students to present cogent and evidence-based argument both in written and spoken language. They wanted more than regurgitation of lecture notes and text book paraphrase. Fundamental to their desired learning outcomes was the notion of *deeper learning*.

7.2.1 *Deeper learning*

One of the prime aims behind this project was to promote *deeper learning* in the students. The unit has been constructed to meet the requirements of the university, faculty and professional organisations. It had defined objectives and espoused promotion of the university's *graduate outcomes.* The subject structure lent itself to dissemination of various philosophies and theories through lectures and tutorial readings. Assessment purported to evaluate students' critical evaluation and understanding in this context. However, neither the tutorial classes nor the tasks themselves afforded support in acquiring the desired critical and evaluative skills. In addition, students saw little relevance in the unit and prepared for the final exam by learning from the prescribed text, despite being warned that it did not give exhaustive coverage of the subject content. Resulting student performance demonstrated only surface learning. Frustrated staff sought a remedy.

Deeper learning encompasses not just evaluation and understanding in the limited context of a single subject, but extension of the learning to future learning, professional practice and even to life in general. Hence, promotion of *deeper learning* (for an extended discussion see Chapter 4, paragraph 4.1.2), should address development of *meta-cognition* (for an extended discussion see Chapter 4, paragraph 4.1.1). Carnell and Lodge (in Klenowski, Askew and Carnell, 2006) recommended that learning should engage students in a cycle of active engagement, review of content, learning about content and application of content. Watkins also specified the importance of incorporating an overarching tier, requiring engagement of students in a further cycle of review of their learning, learning about learning and application to future learning. (Klenowski, Askew and Carnell, 2006)

Lugenbehl (2003) described *deeper learning* as the type of learning that transforms the learner. He claimed it transcended the acquisition of concepts and skills and included a recognition of relevance in the learner's personal and professional life. He proposed four paths to instilling this approach in students:

- Ensuring that students understand why the knowledge is important to them;
- Modelling the behaviours desired for the students;
- Rewarding students as they demonstrate the behaviours;
- Highlight/generate opportunities for students to use the learning outside the classroom.

In designing the learning experiences, he advised instructors to decide what they hope that their students will take away from their subject and maintain focus on imparting its achievement. (Lugenbehl, 2003)

Students extrinsically motivated by the short term goals of 'passing the course', meeting course requirements and even grades, commit new knowledge to short term memory only. They predominantly make connections within the context of a task or, at most, the subject. They rarely uncover external relevance. (Klenowski, Askew and Carnell, 2006) Expository teaching styles that aim to impart clearly specified knowledge in a *one way* transfer to students appear to promote such *surface learning* (Entwistle et al., 2000; Biggs, 1999; Trigwell et al., 1999) and do not address the higher order thinking sought by the teaching staff of this subject. Alternatively, *deeper learning* requires students to look beyond basic facts and processes to determine patterns and relationships external to a task, and identify relevance to their lives. This type of learning is usually exhibited by intrinsically motivated students. To foster this type of learning, teaching should relinquish control of the educational process and facilitate the active engagement of students in a *shared journey* of learning. Instruction should model expected thinking and tasks should encourage students to emulate it. (Johnstone, 2005)

7.2.2 Defining the subject learning outcomes

As this subject was outside the field of expertise of the researcher, prior assessment tasks were used to precipitate intensive discussion with participating staff. The aim was to assist them in *articulating* the learning that they hoped to observe in their students. There was a great deal of commonality in their wish lists, enabling the generation of a clearly defined set of learning objectives for the subject. Relevant *graduate attributes* were also selected from those globally specified by the University.

The transformed learning objectives are illustrated in Table 7.1. Student perception of achievement and the average marks achieved in the corresponding exam questions are reported later in this chapter. These objectives were used to design an aligned and supportive pedagogy. By defining the objectives behaviourally, they also formed foci for designing the assessment. Again the revised taxonomy of Bloom (Anderson and Krathwohl, 2001) was used to analyse assessment tasks and align them with the defined learning outcomes.

Table 7.1 Redefinition of the learning objectives 2005

"… to have developed an appreciation of accounting theory including the types of theories and different research methods, evaluating the underlying assumptions, objectives, paradigms, and logic and knowledge claims of each" [1]
"… to be able to identify the essential elements in the various accounting models available" [1]
"… to be able to demonstrate an understanding of current value income measurement models of accounting developed to date including the impact of inflation on income and value determination" [1]
"… to be able to discuss and evaluate the strengths and weaknesses of traditional accounting systems" [2]
"… to develop an improved ability to critically appraise or evaluate ideas" [2]
"… to be able to appreciate the anthropological view of accounting" [2]
"… to be able to make an effective presentation" [3]
"… to be able to work more effectively in small groups" [3]

Associated university graduate attributes:

[1] demonstrates comprehensive, coherent and connected knowledge

[2] applies knowledge through intellectual inquiry in professional or applied contexts

[3] commands multiple skills and literacies to enable lifelong learning

Source: The UWS Graduate Attributes, 2007

7.3 A pedagogical facelift

The researcher chose the tutorial classes as the main arena for implementing pedagogical changes which might promote the desired deeper learning. Here the class size was more accommodating of an interactive style of teaching/learning that might engage students in collaborative and reflective appraisal of the skills they needed to develop. Most of the strategies selected for implementation, have been discussed more fully in Chapters 3 and 4.

7.3.1 Scaffolding learning: rejuvenating tutorial classes

In this project, an attempt has been made to maintain and extend the pre-existing element of self-direction in class preparation and to include collaborative learning opportunities.

For this project, the tutorial tasks were to be specifically designed to allow:

- *deeper learning*;
- *active engagement* of the students in *meaningful tasks*;
- *collaborative* construction of *group* responses;
- presentation of group responses;
- individual *portfolio* entries detailing the group responses;
- *peer evaluation* of the group presentations using *marking criteria*;
- personal and group notes; and
- reflections on learning.

Tutorial tasks were placed on the web, and students were still expected to make preparatory notes based on the given resources. During tutorial classes, students formed into small groups (3-5 in each group) and each group collaborated to prepare a presentation based on the focus question, using their personally prepared notes and the prescribed scholarly articles to support their discussions/arguments.

In addition, the students were given copies of the *marking rubric* that was to be applied to their own essays and presentations for assessment. The *rubrics* were structured around the behaviourally framed objectives for the subject assessment. The students were expected to use them to *guide* and *organise* their responses. The intention was to focus students on the learning that they were expected to demonstrate. Greater exposure to these criteria before assessment was expected to focus students.

Each week, two to three groups would present to the class their group's response to the tutorial task. These presentations were *peer evaluated* using the *marking rubric*. These peer evaluations were made available to the presenting groups for their reflective comments. Students were expected to justify their comments with reference to the criteria.

Students aggregated their personal notes, the group notes and copies of peer evaluations in a student *workbook.* The notion of the *workbook* was extended to include important aspects of a *portfolio of learning* and students were to be permitted to use it as a supportive resource in their final exam. All commentaries were to be included in their portfolio, together with a short reflection on their learning. The student *portfolios* were marked for the presenting groups the week following their presentations.

By *practising* the construction and presentation of argument and *observation* of the other groups' achievements against the criteria during the peer evaluations, students were focused on the desired learning through the agency of the marking rubric. They became very familiar with *type of thinking* they were expected to demonstrate.

7.3.2 Promoting learning: Active engagement in meaningful tasks

This subject did not present students with conclusions on existing professional issues, research and theory, but rather highlighted those issues and concepts and engaged students in the debate. Tutorial and assessment tasks required students to either work through practical, discipline-related problems, or to reflectively discuss professional issues. Scholarly articles related to the relevant concepts and theories provided both the stimulation and evidence-base for reflective argument. (Johnston, 2005; Lugenbehl, 2003; Nieweg, 2000; Avery, 1999; Kolb, 1984)

7.3.3 Promoting learning: Collaboration

> The ability to collaboratively create, to write, and manage tasks and projects is becoming increasingly important in the business world. (Pfaff and Huddlestone, 2003, p.37).

Not surprisingly then, teamwork is widely espoused as a desirable graduate attribute in most tertiary institutions. In a collaborative learning environment, students construct meaning not only through their personal interaction with the subject content, but also:

- interaction with other learners;
- observation of the learning of others and
- reflection on both their own learning and that of others.

Constructing meaning in this way broadens perspective and affords opportunities to *deepen* student learning. It can also promote *meta-learning* through the call to the higher order skills of critical analysis, evaluation and judgement. (Kuh, 2003; Pfaff and Huddlestone, 2003; Hernandez, 2002; Mohammed, Klimoski and Rentsch, 2000; Livingstone and Lynch, 2000) Collaborative learning has been discussed in more detail in Chapter 4.

7.3.4 Promoting the learning: peer evaluation

Involvement of students in the assessment of peers may present further opportunities for development of the desired knowledge and skills. In addition, effective university assessment should promote learning but also be sustainable in terms of cost, time and effort (Juwah, 2003).

Juwah claimed that:

> Peer assessment is an interactive and dynamic process that involves learners in assessing, critiquing and making value judgement on the quality and standard of other learners, and providing feedback to peers to enable them to enhance performance. (Juwah, 2003, p.2)

He acknowledged the importance of collaborative learning in a constructivist learning framework and highlighted the need for students to appreciate the advantages to be gained from peer evaluations. To this end he recommended the involvement of students in the setting of criteria for peer evaluation, some degree of training in application of those criteria and some tutor checks on the quality of the marking (Juwah, 2003). Student engagement in the marking process should foster a reflective understanding of the desired learning outcomes and also facilitate *meta-learning*. (Juwah, 2003; Liu, Zhuo and Yuan, 2004)

7.3.5 Promoting the learning: portfolios

A *portfolio* is generally regarded as a collection of an individual's work that usually includes some component of reflection and documents both the development of generic skills and academic achievement. For portfolios to be effective, students need to be aware of their purpose and construction. (Klenowski et al., 2006; Liu et al., 2004; McElwee and Evans, 1992) In order to achieve its potential as a learning tool, the *portfolio* should at least include

> … a range of small tasks throughout the learning programme to ensure that participants are actively engaged in learning activities that can culminate in the final assessment'. (Brown, 2003, p.7)

The reflective comments facilitate opportunities for *meta-learning*. Klenowski et al. (2006) participated in an action research project as part of reflective practice in their professional development as tutors and founded from a co-constructivist perspective. The focus of their case studies was the use of portfolios in both formative and summative assessment. Three cases studies tracked different emphases in the use of portfolios for assessment of learning:

- professional development record: a student selected reflective learning record which was also used for summative assessment;
- learning portfolio: encompassed all aspects of formative assessment including peer-assessment, self-assessment feedback and was designed to prepare students for summative assessment;
- learning record: focused on the development of *meta-learning* through participation in critical and evaluative discussion of content and concepts, together with reflective comment on the personal learning journey. (Klenowski et al., 2006)

Much of these notions underpin the use of portfolios in this subject. The work of Klenowski et al. on portfolios has highlighted the following observations which are highly relevant in this case study:

- they foster personal ownership of learning;

- their purpose defines the internal structure necessary to maintain the continuity of purpose;
- they present opportunities for self reflective learning particularly in a professional context; and
- they create a learning tool which promotes scholarship and critical analysis. (Klenowski et al., 2006)

Liu et al. (2004) have also added that ownership also promotes personal responsibility for learning.

7.4 Refining the assessment

In order to address the perceived deficits in student learning, students needed to be made aware of the learning they were expected to demonstrate, and their learning shortfalls readily identified as a result of formative assessment. To redress their shortfalls, identification of pathways toward remediation were required.

7.4.1 The assessment structure

There was no perceived need for change in the assessment structure. There was some need, however, for marking of the assessment to accurately reflect and target the desired learning and to lend greater transparency to the marking process. As multiple markers applied still more multiple interpretations across multiple campuses, students were not convinced that marking was fair and equitable. The issues of *fairness* and *accountability* are inextricably bound. Hence, addressing one issue also addresses the other. Set marking criteria aligned with all aspects of the teaching/learning framework provide teacher with

- a standard for marking and hence some measure of *comparability* with other markers and *accountability* to students;
- indicators of student learning achievements and deficits; and
- a mechanism for alignment of teaching with assessment.

7.4.2 The marking criteria

The marking criteria were inspired by two ideas:

- The concept of Ausubel's *organiser* potentially affords an organisational structure for student construction of supported argument (Ivie, 1998; Ausubel, 1978; Ausubel, 1960) and
- More recent research on scaffolding has highlighted the possibility of facilitating development of student learning from identification of isolated facts and ideas to the more complex tasks revolving around evidence-based argument (Bliss et al., 1996; Rosenshine et al., 1992).

More detailed discussion of *organisers* and *scaffolding* can be found in Chapter 4. The intent in this study was to provide students with a conceptual framework

> "…for construction of a logical argument involving critical evaluation of evidence. The intention …(was) to provide structures which met the cognitive demands targeted by the task objectives. Thus students were introduced to the *organiser* as an *outline.*" (Morris and Puttee, 2006b, p.139)

Again the challenge was to define the desired learning in terms of observable behaviours specified at the most fundamental level possible. If this could be achieved then:

- opportunities for acquiring the relevant knowledge and skills could be provided;
- assessment could be devised to recognise their achievement;
- teaching could be focused towards development of the requisite concepts and skills;
- achievement of these concepts and skills could be 'checked off' as they were demonstrated; and
- appropriate, detailed and timely feedback could be given.

A marking guide incorporating clearly defined marking criteria, carefully aligned at a fundamental level with the learning outcomes, might fulfil all of these roles. Such a marking guide was prepared to be distributed with the Presentation and Essay assessment tasks and to form the basis for evaluating them. It was also to be used to grade the essays in the final exam. Thus, from the outset, students were to be made aware of the criteria against which their work was to be judged and they would be able use them to help construct their responses. When returned, the *checked* criteria would afford valuable feedback to both students and teachers.

However, it was the development and implementation of this device which proved most problematic for the researcher. Intense discussions preceded their construction, but at the conclusion of each session, some teachers were reluctant to commit to applying a final grade to a student's work as a summary of achievement based upon the criteria. There appeared to be a presumption that the criteria could never be exhaustive and that hence markers should not be bound by them.

The position of the researcher has been that assessment provides the opportunity to detect achievement of set learning objectives. Whilst this does not preclude student demonstration of learning beyond that specified (indeed a highly desirable outcome in any educational environment), such demonstration should neither detract from nor augment achievement recognition within the confines of the assessment regimen. In this study, the criteria focused solely on recognition of *defined* achievement and vindicated grades awarded.

For students, a well-constructed set of criteria provides them with defined learning goals, structures to facilitate their responses and useful feedback on achievement. It can also inspire a perception of *fairness* in assessment. These benefits ultimately became evident in student survey responses.

Working criteria were ultimately devised (see Table 7.2). Since the two main assessment tasks required demonstration of similar knowledge and skills, albeit in different modes (oral presentations and essays), the marking criteria were deliberately closely aligned. However there remained scepticism among the markers about their workability and their potential to exacerbate the arduous task of marking. One member of the team commented that although he could see the benefit of the rubric, he reserved the right to give a *mark independent of the rubric because sometimes he might perceive that a student had performed either better or worse than the scoring showed.* Efforts were made to minimise such a possibility, by returning to more detail in the criteria specification.

Table 7.2: Essay Marking Criteria for Accounting Philosophies and Theories

Please see print copy for Table 7.2

[1] The above outcomes do not all rank equally, but your overall level of achievement will be reflected in the final mark.

(Morris and Puttee, 2006b, p.146)

It was not surprising when these criticisms again arose during the marking of the final exams and teachers partially reverted to their earlier practices. They moved away from the achieved/in progress/not achieved selections to ranges of marks for each of the criteria. Defining criteria/exemplars for the new ranges were not specified. Marks were then *subjectively* awarded from the newly specified ranges. Markers claimed that this gave them more flexibility (perhaps not always a *fair* practice) without departing radically from the criteria to which the students were accustomed. Since students rarely challenge marks in examinations, there was no mechanism for evaluating the validity or reliability of individual decisions and hence their *fairness.*

7.5 Evaluation: seeking evidence

Deeper learning is more likely to result when students regard:

- the subject's presentation favourably;
- the learning goals to have been clearly defined; and
- the assessment as fairly representing achievement of the defined learning objectives and not merely regurgitation of learned facts (Morris and Puttee, 2006b; Barrie et al., 2005; Johnstone, 2004).

Hence evaluation of the success of the strategies implemented in promoting *deeper learning* has sought corroboration from multiple sources encompassing: student perceptions; student assessment, peer review and reflective practice.

7.5.1 Student surveys

An extensive student survey (see Appendix 7.3) was constructed for the purpose of evaluation, and it was distributed to students in tutorials one to two weeks before the close of the session. Students were surveyed for their:

- reported attendance patterns;
- perceptions of the importance of subject presentation to learning;
- perceptions of overall and topic learning;
- perceptions of outcome achievement; and
- group work experience;

7.5.2 Student assessment

Departmental records have provided detail of the results of student assessment. This includes student marks and grades for essays, presentations and final examinations. An annotated copy of the final exam has been included in Appendix 7.4. Analysis has included correlation studies of all assessment tasks and deconstruction of the final exam questions using the revised taxonomy of Bloom (Anderson and Krathwohl, 2001).

7.5.2 Peer review

Throughout the course of the study, peer discussion, conferences, refereed publication of the researcher's work, and reflective practice have provided valuable input for adaptations prior to the implementation and throughout its evaluation.

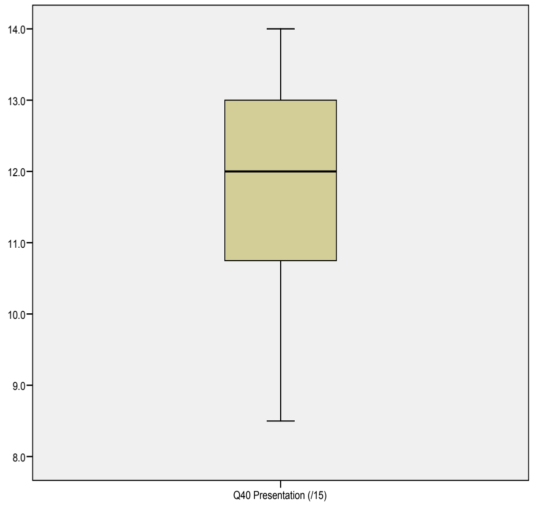
7.6 Checking learning: student survey

The surveys were manually distributed in the final tutorial classes and collected at the end of the session but as assessment tasks were due, many students were absent. A little over 67% (150 out of 223) of students responded to the survey. As this case study was tracking improvement in student learning, it was important to detect any existing bias in student survey responses, particularly in terms of level of academic achievement.

Figure 7.1 demonstrates the comparison between the reported presentation marks and assessment data for the same task. There are some dissimilarities evident: the reported marks show a slight negative skew and the assessment data reveals a slight positive skew; mean and median are slightly higher in the reported marks; and the reported marks show slightly greater variation. The means are significantly different and a comparison of the distribution of the two sets of marks demonstrates a slight bias towards students with higher achievement. This might be expected as more diligent students are generally more cooperative in reflecting on their learning experiences.

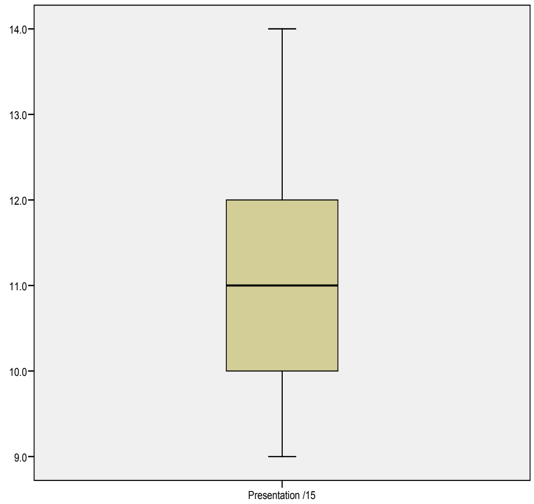
The minimum reported mark may be indicative of poor recall in that no such mark is evident in the assessment file. Although marks reported in the survey varied from those recorded in the assessment file, the purpose of this study was to examine student learning, and the bias towards more successful learners still renders their responses pertinent to this study. Neither were the distributions so different as to lead the researcher to believe that the survey responses did not reflect common attitudes of the students enrolled in this subject.

Reported Presentation marks[1]



		Statistic	Std. Error
Presentation Mark (/15) (N=95)	Mean	11.7	.1
	Median	12.0	
	Std. Deviation	1.4	
	Minimum	8.5	
	Maximum	14.0	

Presentation marks[2]



		Statistic	Std. Error
Presentation Mark (/15) (N=223)	Mean	10.9	.1
	Median	11.0	
	Std. Deviation	1.4	
	Minimum	9.0	
	Maximum	14.0	

Figure 7.1: Comparison of distributions of reported and actual Presentation marks from the assessment data file

[1] Source: student surveys

[2] Source: student assessment data

Appendix 7.1 summarises not only the changes to pedagogy for this case study but also associated student survey responses. Appendix 7.2 records the changes to the assessment regimen in consequence of implementing the teaching/learning framework. The associated mark summaries for each aspect of assessments have also been given.

7.6.1 Student survey: attendance patterns

Throughout this study, although attendance at lectures and tutorial classes was compulsory (students were required to *show cause* if their attendance rate fell below 80%), students generally appeared to find attendance less important for lectures (40% reported attending between 11-13 weeks) than tutorial classes (91% reported attending between 11-13 weeks). The attendance pattern for tutorials was also substantiated by the rankings of student perception of their importance to learning (see Table 7.4). Perhaps the number of requests in general survey comments for release of the notes before lectures reflected student belief that the notes could replace attendance at lectures.

Table 7.3: Aspects of the teaching/learning framework ranked by percentage of students perceiving them as moderately to extremely important to their learning (N=150)

Please see print copy for Table 7.3

[1] Refers to marking criteria/guides which were given with each assessment task and used for peer evaluations.

Source: Morris and Puttee, 2006b, p.141

7.6.2 Student survey: importance of subject presentation to learning

The relative order of the rankings of the different aspects of the subject's presentation (see Table 7.3) afforded few surprises. The smaller group teaching environment and the continuous assessment tasks were highly valued. The lower ranking of the presentation exercises and peer evaluations might be surprising as these underpinned the tutorial classes and 86% of students reported these as moderately to extremely important to their learning.

The high dependency on the role of tutor (84% reporting as important to learning) may reflect the expository teaching style to which the students had become accustomed throughout their degrees. If this is true, however, then it is difficult to understand the lower ranking of the lectures and lecture notes. This might be attributable in part to their late release on the web.

Their value for assessment as a driver for their learning (Biggs, 1999) is reflected in the proportions reporting the essays and presentations as important. Equally their perceived value for the portfolios is possibly attributable to the support afforded by as a resource in and for the final exam.

Student failure to perceive the advantages of the implemented tutorial strategies possibly reflected shortcomings in their implementation. One student explained:

> … spend time discussing articles as group presentations were done before discussion and did not help' (Student survey comment, 2005).

Individual preparation (or the lack of it) did not appear adequate to stimulate group discussion. Some tutors also voiced their concerns that the process of peer evaluation was not treated seriously by students.

This student behaviour has stimulated thought as to how the value of the exercise might be enhanced. Engagement through training in the purpose and application of the criteria might have incited greater commitment (Juwah, 2003). A *teacher directed discussion/development of the criteria* might provide more effective cognitive *scaffolding* for student evaluations (Bliss, Askew and Macrae, 1996). A further possibility might be to require students to include in their portfolio at least one peer evaluation of another student's work together with an appropriate reflective comment or statement of justification. Such strategies would further reinforce student development of critical evaluation skills by:

- promoting more considered application of the criteria by which they themselves are judged and
- focusing their attention on judgement might limit facile checking of the criteria. (Morris and Puttee, 2006b)

7.6.3 Student survey: perception of learning

Only 52% of respondents reported achieving the subject objectives (Table 7.4). This reflected the teachers' perceptions of student belief that the subject was difficult, but was nevertheless disappointing since efforts had been made to increase their confidence. However, consideration must be given to the fact that most of these students were final year students and that their survey comments highlighted a preference for more traditional textbook approaches to teaching and learning. Some comments sought a return to more expository teaching in tutorials:

> … teach as group presentations were a waste of time (Student survey comment, 2005).

Others appreciated the subject structure:

> … open ended and articles facilitated much discussion (Student survey comment, 2005).

The approach used in this session was novel in comparison to their previous experience of accounting subjects.

Table 7.4: Percentage of students responding to overall subject learning (N=150)

Please see print copy for Table 7.4

Source: Morris and Puttee, 2006b, p.142

There was room for extending the number of options for responses to overall learning. A further category *Challenged but achieved many of the objectives* might have captured students who did not believe in their achievement of all of the objectives, but recognised meeting a substantial proportion of them. The options offered no middle ground between *total* or *little* achievement.

7.6.4 Student survey: topic learning

Student survey comments indicated that they were *more comfortable* with a traditional approach of textbook-based content and questions. This contention was supported by student ranking of topic achievement (see Table 7.5). *Agency & Positive Accounting Theory* which encompassed more traditional theories inspired the most confidence in students. This topic is readily found in traditional textbooks, unlike the more intellectually confronting *Critical Theory* (in which the lowest proportion of students reported confidence). For this subject, knowledge of more recent discipline research was provided by scholarly journal readings for relevant topics and was not to be found in more traditional textbooks.

Table 7.5: Confidence in topic learning ranked by percentage of students (N=150)

Please see print copy for Table 7.5

[1] Traditional theories frequently encountered in traditional courses and practice.

[2] Topics for Part B of the final examination: Models: Historical Cost (HC); Continuous Contemporary Accounting (CoCoA)

Source: Morris and Puttee, 2006b, p.142

Nevertheless the proportions of students reporting confidence in each topic (including *Critical Theory*) exceeded the proportion reporting achievement of the subject objectives (Table7.4). Perhaps their overall perception of subject difficulty has been clouded by the more challenging topics.

7.6.5 Student survey: perceptions of achievement of subject outcomes

Teachers were not surprised by the proportions reporting achievement of subject outcomes (see Table 7.6). They believed that the students generally lacked the ability to integrate theory and practice. Students too had this perception. They have voiced their perceived lack of achievement of the outcome: *Appreciate the anthropological view of accounting* (see Table 7.6), an objective requiring this skill. Equally, many were dissatisfied with their ability to *Critically appraise & evaluate ideas* and their exam performance against this objective appeared to support this belief (see Table 7.10). Exam achievement too was significantly higher when specifically relevant evidence could be identified through classroom experience, for example question 5 (see note 5 in Tables 7.6 and Table 7.10), but diminished in more general questions for example question 3 (again refer to the notes in Table 7.6 and Table 7.10).

Students were, however, content with their ability to work collaboratively and make effective presentations. The percentage reporting their ability to *Make an effective presentation* reflected their confidence in their ability to discuss issues and ideas *orally* and to effectively use technology to support their discussion. This outcome addressed perception of achievement of two of the graduate attributes targeted by teaching and learning in this subject:

1. Communication skills: communicates effectively through reading, listening, speaking and writing in diverse context;
2. Technology literacy: applies communication and other technologies effectively in personal and professional learning. (The UWS Graduate Attributes, 2007)

Table 7.6: Perceived competence in objective achievement ranked by percentage of students (N=150)

Outcome	Challenged but moderately competent (%)	Extremely competent (%)	Total competent (%)
Work in small groups	44.0	32.7	76.7
Discuss & evaluate strengths & weaknesses in accounting systems [2, 3]	58.0	13.3	71.3
Make an effective presentation	50.7	20.0	70.7
Appreciate different theories of accounting [1, 2, 3]	54.7	7.3	62.0
Demonstrate understanding of current value income measurement models [4]	47.3	12.7	60.0
Identify key elements in accounting models [4]	54.0	6.0	60.0
Critically appraise & evaluate ideas [3]	47.3	8.0	55.3
Appreciate the anthropological view of accounting [5]	38.0	12.0	50.0

[1] Broadly assessed in Exam essay question 1, selected by 54% of students with a mean of 9.4/20

[2] Broadly assessed in Exam essay question 2, selected by 47% of students with a mean of 9.7/20

[3] Possibly assessed in Exam essay question 3, selected by 41% of students with a mean of 8.5/20

[4] Specifically assessed in Exam essay question 4, selected by 55% of students with a mean of 8.7/20

[5] Specifically assessed in Exam essay question 5, selected by 94% of students with a mean of 20.7/40

(Adapted from Morris & Puttee, 2006b, p. 143)

7.6.6 Student survey: experience of group work

> 74% of students reported a positive experience of group work although many expressed preference for no group involvement (Morris & Puttee, 2006b, p.141)

The responses have been tabulated in Table 7.7. 76.7% of students regarded themselves as moderately to extremely competent in working in small groups, the highest proportion reporting achievement of any outcome (see Table 7.6). Nevertheless, some student comments indicated dissatisfaction with working in groups, claiming inequity in individual commitment while others complained that they would have preferred

> …more discussion and analysis… (Student survey comment, 2005)

Table 7.7: Percentage of students reporting experience of group work (N=150)

Response	Percentage of students
Worked well and learned	39.3
Prefer not but still a positive experience	34.7
Prefer not and unequal effort by participants	11.3
Prefer not and no one learnt anything	2.0

(Morris and Puttee, 2006b, p. 142)

7.7 Assessment

7.7.1 Formative assessment

Formative assessment included:

- an essay: individually composed and submitted toward the end of the session;
- a presentation: group work based on the accumulated research of the members' essays and
- a participation mark: based on *pro-active* attendance at tutorials and completion of the portfolio.

The group presentations were based on the ideas developed in the individually submitted and written essays, and as both the presentation and the essay were marked against similar criteria (and because of student familiarity with these criteria), they might be expected to be significantly correlated. The correlation was significant (p value <0.026, N=222) but weak (r=0.149). There are three possible explanations for this:

1. Different skills were involved in the two tasks. Student perceptions of their ability to communicate critical evaluation orally are amply evident. Not so their confidence in their written skills!
2. Students perhaps feel more confident as group members than as individuals; and
3. The groups' work was aggregated for the presentation whereas the essays were independently constructed.

It should also be noted that the teamwork skills involved in the group presentations were not targeted for assessment. Possibly this could be considered in future implementations.

The paired differences for all pairs of continuous assessment tasks (scaled out of 20) were statistically significant ($p<0.05$). The mean Presentation mark was proportionally higher than both the mean essay mark and the Participation marks and the essay marks displayed the greatest variability.

Table 7.8: Descriptive statistics for the continuous assessment

	N	Minimum	Maximum	Mean	Std. Deviation
Presentation /15	222	9.0	14.0	10.9	1.4
Essay /20	221	5.3	18.9	12.3	2.9
Participation /10	222	4.0	9.0	6.3	1.2
Assessment Total /45	222	14.5	38.8	29.5	4.1

7.7.2 Summative assessment

The pass rate of 88.3% might appear at odds with the students' personal perception of their learning. Student attitudes probably derived more from their perception of the subject's difficulty than confidence in their learning. Around 52% of students report that they believe that they had achieved the subject's objectives (Table 7.4). This is not too different from the 47% of students who achieved more than 50% in the final exam. However the overall level of achievement was moderated by the 95% who were awarded at least 50% in the continuous assessment and hence the resultant higher pass rate.

The final exam contained a compulsory question (Part B) which involved the numerical solution of an accounting problem and was related to the traditional models with which many students were comfortable. For the essay responses, students were required to select three questions from a set of five which related to the theories developed through the lectures.

A review of the learning objectives (Table 7.1) and the final assessment exposed a complex pattern of alignment. A copy of the final exam, annotated by the primary teacher/marker, has been included in Appendix 7.4. The apparently more difficult questions were aligned with multiple objectives (see the notes in Table 7.10). The need for caution in the uneasy teaching path between exposing content and promoting skill development was strongly highlighted. Lecture notes provided illustrative content and structure readily adapted for some exam questions:

- question 2:

  > '…practically based using specific theories …we had discussed a number of traditional theories relating to this issue' (teacher/ marker evaluation comments, Appendix 7.4);

- question 5:

  > 'More of a critical thinking question but linked to the text and specific authors in this area' (teacher/ marker evaluation comments, Appendix 7.4).

Other questions, for example Question 3, proved more challenging:

> '…about current practice but from an interpretive view and not so easily quoted…' (teacher/ marker evaluation comments Appendix 7.4)

proved more challenging.

Lugenbehl (2003) suggested that there should be a cognitive match of teaching style with the desired learning outcomes. If the teaching focuses on *exposition* of *content* and does not *model* the higher order *cognitive skill* of *critical evaluation*, then students may have difficulty demonstrating such skills in assessment.

Consideration could be given to reformulating the objectives in more generic terms (that is eliminating reference to discipline content) and using the defined weekly content as the vehicle for promoting their achievement. The current hybrid of skill and content in the subject learning objectives appears to have presented difficulties with alignment of teaching, learning and assessment.

Table 7.9: Pearson's correlation coefficient (r) for student assessment with their final exam marks

	Participation /10	Essay /20	Presentation /15	Assessment Total /45
Exam /100 P value	0.28 < 0.001	0.27 < 0.001	0.09 .187	0..32 < 0.001

It might be expected that effective formative assessment would be related to the final exam marks (Table 7.9). The Participation marks (r=0.3) and the essays (r=0.3) were significantly positively correlated with the final Exam mark but the Presentations were not. All correlations were however weak.

Table 7.10: Descriptive statistics for exam questions (N=223).
All exam questions reported out of 20 for comparison

	Percentage selecting question	Minimum	Maximum	Mean	Standard. Deviation
Exam Q1 /20[1]	54	2.0	16.0	9.4	3.7
Exam Q2 /20[2]	47	2.0	15.0	9.7	2.9
Exam Q3 /20[3]	41	1.0	15.0	8.5	3.2
Exam Q4 /20[4]	55	3.0	15.0	8.7	2.7
Exam Q5 /20[5]	94	2.0	16.0	10.2	2.5
Part B[6]	100	.0	18.0	10.3	3.5
Exam mark /55	223	.00	41.50	26.4	5.4

[1] Generic research question with detail expounded in the text book and in lecture notes. It required an expository and structured response.

[2] Question requiring a comparative evaluation of *justification of current practice versus development* of better systems with reference to readings.

[3] A general, interpretive question requiring *discussion of the reliability of a theory.* Source detail more dispersed throughout notes and more difficult to identify.

[4] Although this question referred to a specific accounting model (also relevant in Part B) it required *an explanation of the theory* and the wording was complex (see appendix 8.3).

[5] Although this question appeared to require critical thinking for *discussion*, again it was a text book based question linked to specific authors. Teachers perceived this as the easiest option.

[6] Topics: Models: Historical Cost; Continuous Contemporary Accounting (structured response to a practical problem but time consuming)

Student perceptions of topic learning (Table 7.5) align with teacher perceptions and indeed have been further supported by their performance in the final exam (refer to Table 7.10). Their selections (3 from 5) from the five essay questions provided illustration.

94% of students selected Question 5. It ostensibly required critical thinking for *discussion of cultural relevance in accounting*. However, it was a text book-based question linked to specific authors and had been covered comprehensively in lectures and class work. The mean mark for this question was only surpassed by that for Part B, although marks for this question showed more variability than any other. Teachers too perceived Question 5 as the easiest option. It aligned with a single *content-based* subject objective, further evidence of student preference for content-based learning which is associated with *surface learning*. The alignment of subject objectives and exam questions has been summarised in the notes to Table 7.11.

The wording of Question 4 was complex, and only 55% of students selected it. Although it referred to a specific accounting model (also relevant in the compulsory Part B question) it required *an explanation of the theory.* Students who selected this question did not generally respond well (mean mark=8.7/20).

Question 1 was a generic research question with detail expounded in the text book and in lecture notes and hence it similarly required an expository and structured response. It was selected by 54% of students and marks for this question showed the greatest variability.

Question 2 (selected by 47% of students) was also a demanding question requiring extraction of evidence from several readings, although several traditional theories had been discussed in class in this context. The mean mark for this question was 9.7/20

The least selected question was Question 3, which was identified as the most cognitively taxing question by teachers. It was a general interpretive question requiring *discussion of the reliability of a theory* and aligned with multiple subject objectives. Source detail was more dispersed throughout notes and more difficult to identify. Question 3 demonstrated the lowest mean mark (8.5/20)

An analysis of variance indicated a significant difference (α<0.05) between the means for the five questions and *ad hoc* pair-wise tests specifically identified Question 3 as the lowest mean with significant differences detected between questions 2 and 3, questions 3 and 5, and questions 4 and 5 (p values<0.03). This introduces no further complication to the above discussion but may reflect the question groupings selected by the students.

It was surprising to find no significant correlation between the exam essays (except for the weak, but significant (p=0.05) relationship between the most commonly selected Question 5 and Question 2 (r=0.2). There was also no significant correlation between any essay question and the responses to Part B! However this had also been evidenced in prior sessions. Teaching staff have observed that because Part B demands the more traditional and structured response which students prefer, they often spend far too much time completing this question to the detriment of their essays.

7.8 Reflective discussion: impacting on student learning

For this case study, the researcher was not strategically placed in the classroom (fundamentally because of her lack of expertise in the accounting discipline) and consequently was not always aware of changes instigated by the teachers throughout the implementation. There were unanticipated diversions from the intended pedagogy.

Since the aim was to develop a teaching/learning framework that developed *deeper learning* for this discipline, it was relevant to reflect upon both student learning and any relevant impacts upon such learning.

7.8.1 *Deeper learning*

Plotting an instructive path encompassing delivery of content and development of skills is fraught with complexity. If the lecturing focuses too closely on developing content/discipline knowledge, then lectures may fail to model the critical and evaluative skills this subject has sought to inspire. The critical and evaluative process of evidence gathering can become lost in the transmission of the structure and detail of the evidence itself. If the aim is to enable acquisition of both knowledge and skills, these need to be individually specified in the learning objectives for the subject. Assessment then should encompass a marriage of these two dimensions of learning (at least – as the affective domain also needs representation).

To optimise *deeper learning*, Lugenbehl (2003) suggested it is essential to effectively:

- provide relevance;
- model the thinking;
- reward desired behaviours; and
- develop external relevance.

In this subject, students appeared to understand the relevance of the desired skills but still appeared to experience difficulty identifying and selecting the appropriate discipline content to independently demonstrate achievement of those skills. There remains a need for modelling the desired thinking in lectures, possibly supported by exemplars on the web.

Student responses to Question 5 in the final exam illustrated their facility in supporting a familiar argument with familiar evidence. Their apparent discomfort with Question 3 in the final exam highlighted weaknesses in assembling an unfamiliar argument supported by evidence from multiple sources, even though they had been exposed to both the requisite discipline knowledge and skills. Some lectures had accentuated the discipline content, and others the skill but the most successful lectures were those that had married the two approaches.

To prevent a cognitive mismatch between assessment and teaching, lectures needed to model the thinking (Johnstone, n.d.). Perhaps the marking criteria might be used to facilitate development of the lectures. A suggested approach appears in Figure 7.2.

Figure 7.2: Possible structure for lecture

Focus/statement/question to stimulate argument/discussion relevant to discipline content	
Discipline content structured against the criteria	Criteria
Introduction to lecture:	
	Define the problem
	State the approach to answering that will be used
	Define key concepts and words
	Acknowledge all assumptions
The body of the lecture:	
	Highlight the points/steps in the argument
	Offer alternative opinions
Conclusion	
	Relate conclusion logically to argument steps
	Acknowledge alternatives
	Discriminate between different approaches to the argument discussed and evaluated
Impact:	
	Discuss ramifications of the conclusion
	Acknowledge potential effects of the conclusion
	Make relevant recommendations

Figure 7.1: Possible structure for lecture

With the formative assessment tasks due so close to the end of the session, tutorial class use of the criteria was important in continually appraising students of what was expected of them. Since the peer evaluation exercises were poorly regarded and little effort was exerted in application of the criteria, any improvement in raising their profile would be advantageous in reinforcing appropriate demonstrations of the desired learning throughout the session.

The authentic tasks and current discipline research articles provide external relevance. This could be further reinforced through the lectures and possibly through the inclusion of a short investigative problem.

7.8.2 Collaborative learning

Students generally worked successfully in teams. However many still expressed preference for working alone. In the interest of promoting *deeper learning*, and in particular, *meta-cognition*, the structure of group work could be further enhanced by:

- inclusion of tasks that allow student collaboration on knowledge and concepts while formulating responses to slightly different problems and
- individual peer evaluations of contributions of members of the group to be included with the group task. This would facilitate evaluation of *meta learning*.

7.8.3 Peer evaluation

As both students and tutors report little appreciation of the peer evaluations using the Criteria, there is need to augment general perception of their importance from the outset of the session. One possibility that would also reinforce *meta-learning* would be to require students to include (for assessment) at least one of their peer reviews in their portfolios. This could augment the reflective component of the portfolio, and would demonstrate their ability to practically apply the critical and evaluative skills. They should be required to provide a full justification for their evaluation. Such an approach might increase focus on both the criteria themselves and their appropriate application. In consequence, more timely feedback could be given much earlier in the session.

7.8.4 Student portfolios

For this session the prime motivation for student construction of the portfolios was their use as an *open book* resource in the final exam. However such *extrinsic* factors fail to optimise the portfolio's potential as a tool for learning. More *intrinsic* motivation is generally provided through student *ownership* of the learning journey (Klenowski, 2006; Liu et al., 2004) and further effort should be made to heighten this aspect.

Augmenting the reflective component to include judgement of other student work using the *marking criteria* highlights not only the desired critical and evaluative skills defined therein, but introduces a further dimension through *meta-cognitive* application of the criteria in evaluating other students' work. Assessment of this dimension would still sit comfortably within the assessment regimen, seeking extension of the subject objectives to even more complex cognition. Such an inclusion would thus enhance the *portfolio* as a tool for development of *deeper learning*.

As mentioned in the previous paragraph, assessment of the students' evaluations would also provide more timely feedback on the criteria than that afforded by the essay and presentation which are returned late in the session.

Although the use of the portfolio as a resource for the final exam did provide a degree of motivation, students were also allowed to use the text book as a resource in the spring session following this implementation. Teachers reported a concurrent diminution of effort in portfolio work and an increase in student complaints about the usefulness of the text. Student results for the spring session demonstrated a return to surface learning and a step back for the project!

7.8.5 Active student engagement

Juwah (2003) recommended inclusion of students in the construction of marking criteria or alternatively in training in their use (Juwah, 2003; Liu, Zhuo and Yuan, 2004) A teacher led construction of the criteria might also provide cognitive scaffolding for their use. (Bliss et al., 1996)

7.8.6 Marking criteria

Reflection on the impact of the criteria needed examination from two perspectives: the teacher and the student. The teaching perspective highlights two distinctly positive ramifications of implementing the marking criteria:

1. defining assessment strategies from the outset facilitates development of aligned teaching, learning and assessment;

2. clearly defined criteria facilitated fair and justifiable allocation of grades for assessment. It allowed promoted rigour and reduced the potential for subjective response to student work. It should also be noted that the students were very familiar with these criteria and consequently, teacher expectation may have been higher

However despite the careful formulation of the marking criteria, and an understanding on the part of students that they would be judged against these criteria for all assessment, teachers resorted to modifications for marking the final exam. These changes were driven by:

- time constraints and the number of essays to be marked;
- marker belief in the staff ability to 'judge fairly' for the given ranges of marks; and
- an underlying lack of confidence in the criteria to effectively capture varying levels of student achievement.

There is need for further work here. Staff had used the set criteria extensively, and yet under pressure reverted to previous practice of 'value' judgments in marks ranging from 0-10. An interesting exercise (beyond the scope of this thesis) has presented itself: a correlation study of a re-mark of a sample of the exam scripts with results for this exam.

The problems exhibited by students in aligning responses with the *organiser* might be assuaged by placement of *exemplars* that *model* the critique responses on the web. These could be structured similarly to figure 7.2. Alternatively *directed questions* which seek specific evidentiary responses from relevant articles might also provide modelling.

7.8.7 Alignment through the objectives

Although reformulation of the objectives has provided the prime focus for developments of an aligned and supportive pedagogy, there is a need for use of the clearly defined expectation of learning to be more closely aligned with teaching and assessment in order to *model* thinking and to avoid the cognitive confusion created by differences in the *intent* and *practice* of teaching.

7.9 Conclusion

This study has not provided conclusive proof of the efficacy of the devised teaching/learning framework. Assessment data from previous sessions was not available for comparison and the researcher was dependent upon the collective recollections of the collaborating teachers. Nevertheless several critical issues have been identified and addressing these issues may serve to improve student learning for future implementations.

Strategies targeted for future improvements to the framework include:

- alignment between lectures and assessment to prevent cognitive mismatch;
- early feedback through improvement of the profile of the peer reviews in tutorial classes;
- strengthening collaborative learning through changes in the structure of group work;
- inclusion of a peer review in the assessment regimen;
- engagement of students in the construction of the marking criteria; and
- refinement of the marking criteria to afford appropriate *scaffolding* for student thinking and to facilitate alignment with assessment intent.

Chapter 8
Conclusions

Anon, to sudden silence won,
In fancy they pursue
The dream-child moving through a land
Of wonders wild and new,
In friendly chat with bird or beast –
And half-believe it true.

And ever, as the story drained
The wells of fancy dry,
And faintly strove that weary one
To put the subject by,
“the rest next time – “ “It *is* next time!”
The happy voices cry.

Thus grew the tale of Wonderland:
Thus slowly, one by one,
Its quaint events were hammered out –
And now the tale is done,
And home we steer, a merry crew,
Beneath the setting sun.

Alice! A childish story take,
And with a gentle hand
Lay it where Childhood’s dreams are twined
In Memory’s mystic band,
Like pilgrim’s wither’d wreath of flowers
Pluck’d in a far-off land.

“All in golden afternoon”, by Lewis Carroll

8.0 Revisiting the evidence

In Alice's awakening, reality melds with Wonderland, as the shaken Red Queen becomes the kitten and Alice reflects on whether her journey through the Looking-glass world was *her* dream or that of the Red King. Evidence emerges around her, and her journey mists into the half life of 'Childhood's dreams'. For Alice, her 'tale is done'.

The researcher/teacher too has taken pause in *her* journey, confronted with images of *her* unfolding classroom *realities.* But unlike Carroll's tale, *her* journey is far from over. The teacher's experiences and her understanding of them must serve to inspire further excursions into the teaching and learning of her students rather than to retire peacefully as ghostly images of past practice. Her review of the evidence redirects her to the aims of this study. Her reflection on her journey seeks affirmation of effective practice and identifies further challenges.

Both of the case studies included in this thesis have examined the impact of transparent definition of the learning objectives in focusing teaching, student learning and assessment on the desired learning outcomes. In the STAT131 study (Chapter 6), the learning focused particularly upon the development of *statistical thinking* while in the Accounting Theories and Philosophies study (Chapter 7), teaching and learning targeted *critical and evaluative thinking*. Review of the literature and classroom practice have highlighted similarities between the apparently discipline specific sets of skills underpinning these two learning outcomes.

Motivation for implementing an outcome-based approach to teaching, learning and assessment initially arose from the researcher's high school teaching experience of an outcome based curriculum. She had witnessed its power to align teaching intent with practice and to define expected learning for students. Review of the literature supported these observations. delMas warned of the necessity to

> … not only make our objectives clear, but that we follow through on these objectives by planning instruction to develop these outcomes and assessments that require students to demonstrate their understanding, reasoning and thinking. (delMas, 2002, pp. 1-2)

Biggs claimed that constructive alignment could be secured through behaviourally defined objectives. He observed that

> Because the teaching methods and the assessment tasks now access the same verbs as are in the objectives, the chances are increased that most students will in fact engage with the appropriate learning activities. This is, by definition, a deep approach. (Biggs, 1999, p. 73)

There has been evidence of enhanced student learning throughout this study. However, the reasons for the improvements are complex and inextricably related. Focused teaching and assessment has, however, been driven by simple alignment through the learning objectives for the subjects concerned. The ensuing paragraphs address the research findings for each of the evaluands discussed in Chapter 5. To facilitate the review, focus has been provided by the criteria and indicators of achievement detailed in Table 5.4, with relevant excerpts included within each paragraph.

8.0.1 Focused Teaching

In both case studies, the process involved in selecting and specifying the learning outcomes served to focus teaching to facilitate achievement of those outcomes. Selection of appropriate strategies in construction of the teaching/learning frameworks was driven by the targeted learning. Lectures and learning tasks were framed by learning objectives and student marking guides reflected assessment objectives.

Table 8.1: Evaluation of aligned teaching

2. Teaching	
1. Facets of the subject's presentation support student achievement of the learning objectives	Positive student survey responses on the value of each facet to their learning; Positive attendance and submission rates; Improvements in student marks and grade distributions; Supportive comments and responses from participating teachers and markers.

Source: Table 5.4: A general *comlist* for student learning

In STAT131, teacher, tutor and marker comments recognised the degree of alignment between the teaching and the learning and student survey responses demonstrated their value for the various aspects of the teaching/learning framework. Although many students did not appear to value the agents of alignment (objectives and marking guides), individual student comments identified the focused structure of the subject.

In Accounting Theories and Philosophies, the majority of students (more than 80%) valued most aspects of the teaching/learning framework, but again it was the mechanism for focusing their learning that they did not appear to value (marking criteria, presentation exercises and peer evaluations). One student commented

> … make link between lectures and tutorials more obvious (individual student survey comment, 2005)

The subject teachers also commented on the merit of the *process*, but constructive alignment of the lectures, tutorial tasks and assessment to enable students to make meaningful connections in all topics has yet to be achieved.

Attendance at lectures remained less than 50% in both subjects, but tutorials and laboratory classes, where interactive teaching and collaborative learning drove the tasks, drew in excess of 85% in all sessions. There were very few (less than 4) 'zero fails' (failures to sit the final exam) in any implementation in either subject.

Student marks improved significantly following implementation of the teaching/learning framework (see Figure 6.7: Box plot of final exam marks across all implementations). This proved even more noteworthy when examination of the exam questions also revealed an increase in cognitive demand (see Table: 6.11 and Appendices 6.27 – 6.29).

8.0.2 Enhanced learning

The evidence of student learning has been drawn from the assessment and student perceptions of their own learning. If this evidence is to be valid, then four questions need to be answered:

1. Does the assessment target the desired learning (both knowledge and skills) and do the marking criteria recognise achievement of the objectives and not just recall and repetition of 'pre-processed' thought acquired during learning?

2. Has there been a positive shift in the final exam marks since implementing the strategies?
3. Are students confident in their learning?
4. Is student confidence in their learning justified?

Table 8.2: Evaluation of student learning

1. Student learning	
1. Assessment tasks target and reward the levels of knowledge and skills identified in the learning objectives *	Correlation of deconstructed tasks, solutions and marking criteria using Bloom's Taxonomy
2. Students have demonstrated achievement of the discipline learning objectives	Assessment results/grades; Teacher responses/comments
3. Student have demonstrated confidence in their own achievement	Survey responses/comments
4. Students *know what they know* and *know what they don't know*	Correlation between rankings of topic assessment results and rankings of student topic confidence

Source: Table 5.4: A general *comlist* for student learning

In the STAT131 case study, the revised taxonomy of Bloom (Anderson and Krathwohl, 2001) was used to deconstruct both the questions and the marking of acceptable responses was examined across the observation phase and three of the implementations. As all tasks were written to align with laboratory learning tasks, and objectives had been aligned with the subject learning objectives, the final exams were also so aligned.

Initially, little heed was paid to the marking, but as the study progressed it became increasingly obvious that targeting skills in questions was only one face of assessment. If students were still being rewarded for facile recall and repetition of 'learned' responses, then their focus remained on demonstrating only those lower order skills! In later implementations, the tasks more directly targeted the more complex thinking, with little room for rewarding lower order cognition in its place.

As mentioned in paragraph 8.0.2, the mean marks in the final exam increased after implementation of the revised teaching/learning framework. This was also accompanied by a shift to higher cognitive demand in the exam questions themselves (see Table 6.11). Teacher comments also indicate higher expectations of student learning that appear to have resulted from their more focused approach to teaching.

Student comments also indicated their comfort with learning in this subject. Students were surveyed on their ‘comfort’ with topic learning (see Appendix 6.19) and their perceived progress toward achievement of the ‘graduate qualities’ (see Appendix 6.25). Not surprisingly, some topics induced more confidence than others. Group confidence in student topic learning was correlated with the mean marks in the corresponding topic questions in the final exam. The correlation was significant ($r=0.9$, $p<0.02$). As discussed in paragraph 6.7.1, this evidence cannot support the claim that students *know what they know* and *know what they don’t know* as the data do not refer to individual marks and the corresponding perceived level of confidence. It does indicate that the class tended to receive higher marks for topics where a higher proportion of students indicated confidence. Nevertheless, the teacher/researcher still speculates that a more detailed study of individual students’ levels of achievement might correlate positively with their level of confidence in their own learning.

The evidence in the second case study is not easily isolated. Although tutorial tasks were aligned with the subject objectives, and modelled the expected learning, the tasks in the final exam represented aggregates of outcomes. Neither were topics readily identifiable. However the primary task was to enhance student achievement of more generic outcomes and both student and teacher general comments indicated some improvements in this direction. One student, however, commented that the skills might be more effectively introduced in this fashion in earlier sessions. As this was an initial implementation, results have recently been under review for implementation of improvements.

8.0.3 Aligned assessment

In STAT131, student survey responses and comments confirm a belief in its ‘fairness’. Most students believed that their learning experience prepared them for assessment. There was no survey of student perception of ‘fairness’ in Accounting Theories and Philosophies. Student comments indicated that many had failed to make connections between the tutorial exercises and assessment. Teacher comments expressed disappointment that there was not any noteworthy shift in student achievement. However, the new marking criteria did afford more accountability in allocation of marks, and student marking disputes were more effectively addressed using them.

Table 8.3: Evaluation of aligned assessment

3. Assessment	
1. Assessment is aligned with the subject learning objectives and the student learning experiences	Positive student survey responses; Improved completion rates; Improved submission rates; Improvements in student marks and grade distributions; Comments and responses from participating teachers and markers. Correlation of deconstructed tasks, solutions and marking criteria using Bloom's Taxonomy

Source: Table 5.4: A general *comlist* for student learning

There was no evidence of change in completion or submission rates in Accounting Theories and Philosophies. There was some deterioration of submission rates in 2005 in Stat131 and some discussion on potentially impacting issues were discussed in paragraph 6.7.4. There was no evidence of changes in completion rates.

In STAT131 there were some subtle changes in grade distributions, but the increase in the failure rate from spring, 2004 to spring 2005 has also been discussed in paragraph 6.7.4. Improvements in student marks have been mentioned in the previous paragraphs as have the comparisons of cognitive demand of assessment and the acknowledgement of achievement through application of the marking criteria.

There were no historical records for comparison of marks and grades in Accounting Theories and Philosophies. Deconstruction of the exam questions revealed a complex aggregation of desired learning outcomes (see paragraph 7.7.2) and the marking criteria were 'altered' for marking of the exam questions (see paragraph 7.8.6)

8.0.4 Constructive alignment

In both case studies, final assessment had been used to facilitate definition of desired learning outcomes. Learning tasks were aligned with these outcomes through specification of the relevant objectives. Assessment was modelled on the learning tasks and knowledge and skills aligned through the objectives. Marking criteria were also aligned with subject learning. The aligned teaching/learning/assessment framework for STAT131 is illustrated in Figure 6.1. Teachers and markers in both case studies acknowledged the benefits of alignment of all aspects of that framework. Student comments (a single student in Accounting Theories and Philosophies) in both case studies indicated that they had detected 'structure' in the subjects.

Table 8.4: Constructive alignment

4. Alignment of teaching, learning and assessment	
1. Assessment is aligned with the subject learning objectives and the student learning experiences	Positive comments and responses from participating teachers and markers. Increased cognitive demand revealed by deconstruction of tasks, solutions and marking criteria using Bloom's Taxonomy

Source: Table 5.4: A general *comlist* for student learning

The increased cognitive demand of assessment in the STAT131 case study was discussed in paragraph 6.7.5. Evidence of enhanced student learning derives not only from increases in marks but also the accompanying increase in cognitive demand of the final exam questions. Without comparative results, the researcher has relied on the teacher and student comments as indicators of enhanced learning in the Accounting Philosophies and theories case study. Although the teachers' believed the subject alignment to be improved, this has yet to translate to student confidence and performance.

8.0.5 *Statistical thinking*

Table 8.5: Evaluation of *statistical thinking*

5. Statistical thinking	
1. Assessment is aligned with the subject learning objectives and the student learning experiences	Increased cognitive demand revealed by deconstruction of tasks, solutions and marking criteria using Bloom's Taxonomy Student marks and grades static or improved; Comments and responses from participating teachers and markers.

Source: Table 5.4: A general *comlist* for student learning

The literature has revealed confused and conflicting use of the terms *statistical literacy*, *statistical reasoning* and *statistical thinking*. For the purposes of this study, the researcher/teacher has adopted the following approach proposed by Chance (2002) following review of the terms:

> While literacy can be narrowly viewed as understanding and interpreting statistical information presented, for example, in the media, and reasoning can be narrowly viewed as working through the tools and concepts learned in the course, the statistical thinker is able to move beyond what is taught in the course, to spontaneously question and investigate the issues and data involved in a specific context. (Chance, 2002, p.4)

The teachers became more focused upon the desired thinking, throughout the study and consequently the tasks focused more clearly on the type of thinking required. Student survey comments acknowledged the modelling of the desired thinking in the learning tasks and the formative assessment. Increases in the cognitive demand of assessment task questions have primarily resulted from specifically targeting the skills underpinning *statistical thinking.* Student assessment responses have more closely approximated *expert* thinking in later implementations. Evidence of more complex thinking was detected following implementation of the revised teaching/learning framework.

Feedback on achievement through the checked marking criteria and access to ‘model’ solutions enabled remediation of deficits in knowledge and skills and recording of the desired knowledge and skills in their workbooks before the final exam.

8.0.6 *Critical and evaluative thinking*

Table 8.6: Evaluation of *critical and evaluative thinking*

6. Critical and evaluative thinking	
1. Assessment is aligned with the subject learning objectives and the student learning experiences	Increased cognitive demand revealed by deconstruction of tasks, solutions and marking criteria using Bloom’s Taxonomy Student marks and grades static or improved; Comments and responses from participating teachers and markers.

Source: Table 5.4: A general *comlist* for student learning

Deconstruction of the knowledge types and skills required for *critical* and *evaluative thinking* has revealed marked similarities to those required for *statistical thinking.* It would appear that the higher order skills required by these types of thinking have been contextualised to support discipline and professional requirements. However, it is vaguely comforting to find commonality of desired skills across disciplines, each grounded within the rich tapestry of individual discipline histories and practices.

For this subject, some elements of cognitive mismatch still limit development of the required thinking. Whilst some students have accommodated the interactive learning, most expressed disquiet at the shift from the expository style to which they have become accustomed throughout their degree program. Until this has been addressed in early subjects, the resultant shallow learning may persist.

8.1 Future exploration

An apprenticeship with discipline experts and the cyclic revisions of the teaching/learning framework in the STAT131 case study have enabled the teacher/ researcher to microscopically examine her students' development of *statistical thinking*. Throughout this study, her continual review of the relevant literature has exposed her to a vast arena of enlightened and innovative practice. This continues to inspire her in her efforts to construct a learning environment that supports acquisition of the desired thinking and facilitates her students' paths to *lifelong learning.*

8.1.1 Classifying learning

Throughout this study, the revised taxonomy of Bloom (Anderson and Krathwohl, 2001) has been used to classify student learning. The researcher has become more adept in its application with practice. However, reliable application has proved difficult. Different perceptions of desired behaviours arose between individuals. In this study, in order to maintain reliability, all classifications of student learning have been undertaken by the researcher and confirmed with her peers.

The difficulties encountered in classifying statistical learning inspired the search for a more discipline appropriate taxonomy. However her reading of the work of delMas (2002), and expanded with reference to Rumsey (2002), Garfield (2002), Chance (2002) and Wild and Pfannkuch (1999) has highlighted commonality in his classification of desired behaviours into the three domains of basic literacy, reasoning and thinking and the researcher's categorisation of higher and lower order skills using the revised taxonomy. This has confirmed the researcher's belief in her appropriate classification of the desired student learning behaviours.

However there may still be room for a more explicit classification of the targeted knowledge and skills. Evidence from this research supports the contention that students are generally aware of the depth of their knowledge and understanding and the researcher has formed the belief that this notion might be used to structure a more discipline specific classification. Examination of student responses to tasks designed to elicit the desired knowledge and thinking could hierarchically identify most commonly correct responses through to least often correct responses. Such an examination would then be associated with conceptual deconstructions of the language of the tasks to develop a classification system based upon facility of student interpretation. The researcher suspects that again she might observe a high degree of association with her current approaches. Nevertheless, the undertaking could provide significant affirmation of her current understanding of her students' learning.

8.1.2 Transparent objectives

Comparison of the two case studies has highlighted the similarities in the targeted thinking. However, it would appear that a more effective teaching/learning framework has been implemented in the STAT131 study. This might have derived directly from the lack of revision cycles in the Accounting Theories and Philosophies study or possibly the drastic change from the accustomed expository approach to teaching. Nevertheless, teaching associates of the researcher (of essay-based subjects) have maintained that teaching *thinking* in a structured mathematical discipline is inherently less complex. The researcher has reflected upon this in the light of her involvement in the two studies. She has detected two main difficulties in accepting this comment:

1. If it were true, mathematicians/statisticians might be expected to abound, and to date this happy possibility has failed to eventuate; and
2. The STAT131 study has targeted similar *thinking* and *its* complexity arises from identification of relevant evidence (and hence, elimination of distracters), selection of appropriate procedures and/or models, and understanding the data in context. The problem solving skills required

are not dissimilar to those expected of students in the accounting discipline.

Differences in student comments have been indicative of their inability to identify meaningful structure in the framework for the Accounting Theories and Philosophies study and the researcher believes that this has led, in some measure, to a lack of constructive alignment. It would appear that students need to detect meaningful relevance in the way they are taught and hence the process in this subject must therefore become more overt.

'Upfront' specification of objectives is only the first step in transparency. Lectures also need to specify and model expected behaviours; learning tasks also need to specify and model the desired learning and in addition align with assessment. Peer evaluation can afford an appropriate mechanism for developing a reflective understanding of the desired learning (Juwah, 2003; Liu, Zhuo and Yuan, 2004). Peer evaluation also requires engagement of the same *critical and evaluative thinking* required of students in this unit. Perhaps inclusion of a peer evaluation as an assessment task may provide a very practical focus for student thinking.

Wild and Pfannkuch (1999) claimed

> If statistical thinking is something that we teach rather than simply to be absorbed by osmosis, then we have to give it structure. Structure can stimulate thinking, prevent crucial areas from being overlooked, and provide something to fall back on when you hit a brick wall. (Wild and Pfannkuch, 1999, p. 242)

They discussed the use of 'trigger' questions ('How?' or 'Why?') as

> … they tend to initiate new thinking in certain directions. (Wild and Pfannkuch, 1999, p. 242)

Structured thinking should be similarly modelled in lectures for the accounting subject and initially scaffolded for early learning tasks as has proved successful in the STAT131 study.

8.1.3 Supported learning

Although the STAT131 study has substantiated the effectiveness of the teaching/learning framework, recent classroom experience has highlighted that further learning gains may be achieved. Some potential improvements have been proposed in paragraph 8.1.2.

> If self-awareness and the capacity for self development are key components of the capable person, then it is important to provide opportunities in the curriculum for these to develop. … with the half-life of knowledge steadily dropping and life-long learning becoming a necessity ... (Yorke, 2001, p. 122)

In Chapter 7, the researcher has also identified the potential for peer evaluation to augment *meta-learning.* In more recent revisions of the STAT131 teaching/learning framework, the teacher/researcher has endeavoured to incorporate a cycle of *meta-learning* by encouraging student reflection on the purpose and functioning of the constructive alignment of the learning environment. It is the teacher/researcher's observation that it is the student recognition of connective structures that promotes the development of *meta-cognition*, that in turn impacts upon learning.

Similarly, in Chapter 6 the researcher commented on the developing importance of *learning strategy* tasks in facilitating student recognition of *connections* between concepts and procedures. They also provide summary material that triggers more detailed recollection. The teacher now uses such strategies to facilitate a *review* and *practice* approach to exam preparation, and has anecdotal evidence supports improvements in exam performance. There is a need for an evidence-based evaluation of these approaches in the future.

8.1.4 Assessing learning

Several aspects of an effective assessment regimen have been highlighted in this research. Ramsden (1992) claimed

> From our students' point of view, the assessment always defines the actual curriculum (Ramsden, 1992, p. 187)

This has several implications in the design and application of effective assessment:

- the defined learning should be behaviourally defined;
- learning tasks should model the desired learning;
- defined learning should be assessed (delMas, 2002);
- assessment needs to engage the desired levels of knowledge and cognitive skills; and
- the marking criteria should acknowledge the *quality* of student performances with reference to achievement of the learning objectives, rather than generate numerical measures that emphasise differences between students without such reference (Biggs, 1999).

Students need to perceive relevance in their learning to achieve optimal response. This relevance may arise through discipline requirements and/or through appropriately targeted cognitive challenge in resolving *authentic* problems that potentially offer students opportunities to emulate *expert* practice.

The *summative* expectations defined by the assessment regimen (and more particularly, explicitly addressed by the final exam) should facilitate determination and specification of the learning objectives (Biggs, 1998). *Formative* assessment, then, can then be designed to provide opportunities to check performance against these objectives and provide timely feedback on performance. That feedback, however, should be task specific and directed toward remediation of individual needs if it is to promote learning (Black and Wiliam, 1998).

Personally reflective comment (scaffolded by specified criteria) in student learning portfolios or peer evaluation scaffolded by marking criteria affords the facility for engaging higher order cognition and hence may form a constructive aspect of student assessment.

8.1.5 Evaluating the learning experience

Seeking evidence of improvements in student learning has resulted in a mixed method approach to this study. As might be expected, much of this has centred on assessment, both student results and the assessment tasks themselves. Student attitudes were also surveyed because the teacher has long believed that ‘good’ students are aware of what *they know* and *what they do not know.* However, student comment and survey responses propelled further explorations of student perceptions of their learning, and this project has offered some substantiation of the teacher’s belief.

The researcher’s survey of her students increased in detail as the project developed and she became confident that student responses were considered and substantiated in assessment. The pattern of change in the student evaluation surveys for STAT131 has been recorded in Appendix 6.18. Recent reading has indicated the effective use of *knowledge surveys* to assess changes to student learning (Jordan, 2007.; Wirth and Perkins, 2006; Nuhfer and Knipp, 2003).

In practice, *knowledge surveys* seek honest perceptions of student belief in their ability to respond to tasks addressing the subject learning objective. They do not respond to the tasks, merely select their perceived ability to respond from a scale of responses. The tasks themselves may be classified by cognitive process or knowledge type demand. Students are surveyed before and after instruction and students encouraged to track their progress by comparison. (Nuhfer and Knipp, 2003) Nuhfer and Knipp (2003) report an observable association between perceived confidence and exam performance, but no detail if the level of significance of the association was given. Wirth and Perkins (2006) report a moderate correlation between exam performance and the responses to the *knowledge survey*. The researcher believes that both the literature and her case study indicate that a more detailed evaluation may yield affirmation of this innovative practice.

8.2 Improving learning by improving teaching

In the thesis, the researcher has attempted to critically examine the impact of a constructively aligned curriculum on improving the depth of student learning in two disparate university teaching environments. The researcher has been motivated by extensive classroom practice in mathematics and statistics in high schools and at universities in Australia. Throughout this work she has undertaken wide reading and engaged in discussion with her peers in order to ground her understanding of her professional practice. Some important issues have percolated through her perspective of student learning and have impacted upon her professional beliefs and practices:

- the burgeoning recognition of students as *life-long learners*;
- the associated need for promotion of *meta-cognition* and *deeper learning*; and
- the intricate fabric of demands driving assessment at universities and the implications of the divide between criterion and norm referenced approaches.

Much of the practice implemented in this project has resulted from the teacher/researcher's attempts to address these issues. Yorke (2001) has provided a broad perspective

> The Dearing Report identified communication, numeracy, the use of technology and "learning how to learn" as desired outcomes of all higher education programs. Of these, the last most clearly goes beyond mere technical expertise … (with the) broader construct of "capability" as a(n) more appropriate underpinning for higher education, particularly in the current political context in which life-long learning is being highlighted. (This construct) … blends subject expertise with a range of other skills to encourage the development of people as "effective operators in the world", whether in the field of employment or broader society. (Yorke, 2001, p.120)

"Effective operators" then need to be confronted by relevant and challenging problems that engage students actively in their resolution and hence inspire a *deep* approach to their learning as

> … deeper approaches to learning are related to higher quality learning outcomes. (Trigwell et al., 1999, p.57)

To promote self-regulation, students need clearly defined objectives:

> Clear descriptions of student behavior or examples of behavior that demonstrate an objective provide concrete goals for students. Once the student behavior is described, different instructional experiences that might lead to the

> goal can be imagined. Therefore, … (such descriptions) provide the impetus for instructional design. If assessments are then derived from the instructional experiences, students can form valid expectations of how their understanding will be assessed. (delMas, 2002, p. 3)

Biggs further claimed that in such a system

> All the components of the system address the same agenda and support each other. … I call this network *constructive alignment.* (Biggs, 1999, p. 64)

The evidence from this study provides further affirmation of the positive impact upon student learning of *constructive alignment.*

One of the issues that challenged the researcher/teacher throughout was the marking of student assessment. The subjects of her case studies both used a quantitative approach to marking of student assessment. The teacher/researcher has long recognised the inherent logic of a criterion-based approach to marking of assessment that purports to evaluate student achievement of defined learning objectives. She has, however, worked in a teaching/learning environment that has practiced a quantitative approach (even extending to 'grading against the curve'). Her attempts to resolve this issue have centred on an 'on/off' marking system recognising achievement/non achievement of the desired learning outcomes. The summative nature of an aggregation of such a marking scheme, however, forces the onus of recognition of higher order cognitive achievement to the task design. This then requires meticulous representation of opportunity to demonstrate the required order of knowledge and skills. Ostensibly the tasks become more cognitively challenging, but an aligned curriculum should appropriately prepare the students for such increased demand. This was certainly evident in the STAT131 case study.

Should this then be regarded as 'teaching to the test'? Such an assertion might well be regarded as an affirmation of the teaching, rather than denigration. For as Biggs has claimed:

> (If) … the assessment is deeply criterion-referenced, incorporating the intended curriculum, which should be clearly salient in the perceived assessment demands … you get aligned instruction, where teaching to the test

> is exactly what you want because it is teaching the intended curriculum (Biggs, 1998, p. 108)

The teacher/researcher has come to realise the power of assessment to both align the curriculum and inspire her student learning.

References

A

ABS (2007). Education and Training Experience, Australia, 1997 (cat. no. 6278.0); and unpublished data, 1997 Survey of Education and Training. [Accessed 18/07/07 from http://www.abs.gov.au/AUSSTATS]

Adler, P. A. and Adler, P. (1994). *Observational Techniques*, in Denzin, N. and Lincoln, Y. *Handbook of qualitative research*, (pp. 377-392). Sage Publications: Newbury Park, California.

Alexander, S. and Hedberg, J. ((1994). *Evaluating technology-based learning: which model?* In Beattie, K., McNaught, C. and Wills, S. (Eds.), *Multimedia in higher education: Designing for change in teaching and learning*, Elsevier: Amsterdam, pp. 233-244.

Anderson, L. W., and Krathwohl, D. R., Eds. (2001). *A Taxonomy for Learning, Teaching, and Assessing: A Revision of Bloom's Taxonomy of Educational Objectives.* Addison Wesley Longman Inc: New York. U.S.A.

Anderson, R., Spiro, R. and Anderson, M. (1978). Schemata as Scaffolding for the Representation of Information in Connected Discourse, *American Educational Research Journal* 15(3): 433-440. [Accessed 31/05/07 through JSTOR]

Angelo, T. (1999). 'Doing Academic Development as though we valued learning most: transformative guidelines from research and practice', a paper presented to the HERDSA Annual International Conference, Melbourne, 12-15 July, 1999.[Accessed 12/8/2008 from: http://www.herdsa.org.au/wp-content/uploads/conference/1999/pdf/Angelo.PDF]

Ausubel, D. P. (1978). In Defense of Advance Organizers: A Reply to the Critics, *Review of Educational Research* 48(2): 251-257.

______(1968). *The Psychology of Meaningful Verbal Learning,* New York: Grune & Stratton Inc.

______(1960). The use of advance organizers in the learning and retention of meaningful verbal material, *Journal of Educational Psychology* 51(5): 267-272.

Avery, P. G. (1999). Authentic assessment and instruction*, Social Education* 63(6): 368-373.

B

Bain, J. (1999). Introduction, *Higher Education Research & Development* 18(2): 165-172.

Barrie, S., Ginns, P. and Prosser, M. (2005). Early impact and outcomes of institutionally aligned, student focused learning perspective on teaching quality assurance, *Assessment & Evaluation in Higher Education* 36(6): 641-656.

Beattie, K., McNaught, C. and Wills, S. (Eds.) (1994). *Proceedings of the IFIP TC3/WG3.2 Working Conference on the Design, Implementation and Evaluation of Interactive Multimedia in University Settings, Melbourne, Australia 6-8 July 1994.* Elsevier: Amsterdam.

Bellhouse, D.R. (C 2003). 'Probability and Statistics Ideas in the Classroom – Lessons from History', paper presented to the 2003 Invited Session on "Using the History of Statistics to Improve the Teaching of Statistics", San Francisco, USA, 7/8,2003. [Accessed 19/5/07 from http://www.stat.auckland.ac.nz/~iase/publications/13/Bellhouse.pdf]

Bigge, M. (1976). *Learning Theories for Teachers*, Harper & Rowe, Publishers: New York.

Biggs, J. (1999). What the Student Does: teaching for enhanced learning, *Higher Education Research & Development* 18(1): 57-75.

______(1998). Assessment and Classroom Learning: a role for summative assessment?, *Assessment in Education* 5(1): 103-110 [Accessed 9/10/07 from Proquest]

Biggs, J. B. and K. F. Collis (1982). *Evaluating the Quality of learning The SOLO Taxonomy (Structure of the Observed Learning Outcome),* Academic Press Inc: New York. U.S.A.

_______(1979). *Classroom Examples of Cognitive Development Phenomena: The SOLO Taxonomy*, University of Tasmania: Hobart.

Black, P. (2000). Research and the Development of Educational Assessment, *Oxford Review of Education* 26 (3&4) 407-420 [Accessed 8/7/07 from ProQuest Education Journals.

Black, P. and Wiliam, D. (1998). Assessment and Classroom Learning, *Assessment in Education* 5(1): 7-74 [Accessed 9/10/07 from Proquest]

Bliss, J., Askew, M. and Macrae, S. (1996). Effective Teaching and Learning: scaffolding revisited, *Oxford Review of Education* 22(1): 37-61.

Bloom, B. S., Ed. (1974). *Taxonomy of Educational Objectives Handbook 1: Cognitive Domain*, David McKay Company Inc: New York.

Board of Studies NSW (2001). *Annual Report 2001*, Sydney: Office of the Board of Studies NSW. [Accessed 6/5/08 from http://www.boardofstudies.nsw.edu.au/manuals/pdf_doc/bosannualreport01.pdf

Brew, A. (1997). Changing Assessment Practices: How are we to understand them?, *Higher Education Research & Development,* 16(3): 371-380.

Brown, S. (2003). Assessment that works, *The Newsletter for the Institute for Learning and Teaching in Higher Education,* Summer, 11, 6-7.

Bruner, J. (2000). *Cognitive and constructivist learning theory*, in Deissner and Simmons (ed.), *Sources: Notable selections in educational psychology; Chapter 7, The Act of Discovery*, Dushkin/McGraw-Hill: Guilford, Conn, pp. 216-227.

______ (1971). *Toward a Theory of Instruction*, The Belknap Press of Harvard University Press: Cambridge.

C

Cameron, J., Pierce, D., Banko, K. and Gear, A. (2005). Achievement-Based Rewards and Intrinsic Motivation: A Test of Cognitive Mediators, *Journal of Educational Psychology*, 97(4): p641-655.

Carnell, E. (2007). Conceptions of effective teaching in higher education: extending the boundaries, *Teaching in Higher Education*, 12(1): 25-40. [Accessed 21/06/07 from http://www.informaworld.com/smpp/content~content=a769868624~db=all]

Carroll, L. (2002). Alice's Adventures in Wonderland and Through the Looking-Glass: and What Alice Found There, The Modern Library: New York.

Chance, B. (2002). Components of Statistical Thinking and Implications for Instruction and Assessment, *Journal of Statistics Education*, 10(3). [Accessed 3/7/2008 from http://amstat.org.ezproxy.uow.edu:2048/publications/jse/v10n3/chance.html]

Cobb, G. and Moore, D. (1997). Mathematics, statistics, and teaching, *The American Mathematical Monthly*, 104(9): 801-821. [Accessed from ProQuest 23/1/06]

Cranton, P. and Roy, M. (2003). When the Bottom Falls Out of the Bucket: Toward a holistic perspective on transformative learning, *Journal of Transformative Education* 1(2): 86-98 [Accessed 4/11/04 from http://online.sagepub.com]

Cresswell, J. (2002). *Educational Research: Planning, Conducting, and Evaluating Quantitative and Qualitative Research*, Merrill Prentice-Hall: Upper Saddle River, New Jersey.

Crooks, T. (1988). The Impact of Classroom Evaluation Practices on Students, *Review of Educational Research*, 58(4): 438-481.

D

Daniels, H. (2003). *Vygotsky and Pedagogy*, RoutledgeFalmer: New York.

Dearing, R. (Chairman of the National Committee of Inquiry into Higher Education (NCIHE)) (1997). *Higher Education in the Learning Society*, Hayes: Middlesex. [Accessed through Google on 16/10/07 at: http://www.leeds.ac.uk/educol/ncihe/]

Deissner, R. and Simmons, S. (2000). *Sources: Notable selections in educational psychology*, Dushkin/McGraw-Hill: Guilford, Conn.

De Mulder, E. and Eby, K. (1999). Bridging Troubled Waters: Learning Communities for the 21st Century, *American Behavioural Scientist* 42(5): 892-901.

delMas, R. (2002). Statistical Literacy, Reasoning and Learning: A Commentary, *Journal of Statistics Education*, 10(3). [Accessed 3/7/2008 from http://amstat.org.ezproxy.uow.edu:2048/publications/jse/v10n3/delMas.html]

Denzin, N. and Lincoln, Y. (2003). *Collecting and Interpreting Qualitative Materials*, Sage Publications: Newbury Park, California.

____ (1994) *Handbook of qualitative research* (1st ed), Sage Publications: Newbury Park, California.

Driscoll, M. (2000). *Psychology of Learning for Instruction*, Allyn & Bacon: Needham Heights.

E

Entwistle, N., Skinner, D., Entwistle, D. and Orr, S. (2000). Conceptions and Beliefs About "Good Teaching": an integration of contrasting research areas, *Higher Education Research and Development* 19(1): 5-26.

F

Fontana, A. and Frey, J. (1994). Interviewing: The Art of Science, in Denzin, N. and Lincoln, Y. *Handbook of qualitative research*, (pp. 361-376). Sage Publications: Newbury Park, California.

G

Gal, I. and Garfield, J. (eds.) (1997). *The Assessment Challenge in Statistics Education*, IOS Press [Accessed 19/5/07 from: www.stat.auckland.ac.nz/~iase/publications/assessbk/]

Garfield, J. (2002). The Challenge of Developing Statistical Reasoning, *Journal of Statistics Education*, 10(3). [Accessed 3/7/2008 from http://amstat.org.ezproxy.uow.edu:2048/publications/jse/v10n3/garfield.html]

Gomm, R. (2004). *Social Research Methodology*, Palgrave MacMillan: New York.

Gredler, M. (2001). *Learning and Instruction: theory into practice*, Merrill Prentice Hall: Upper Saddle River, New Jersey.

Green, J. (2003). *Understanding Social Programs Through Evaluation*, in Denzin, N. and Lincoln, Y. *Collecting and Interpreting Qualitative Materials*, (pp. 590-618). Sage Publications: Newbury Park, California.

Guba, E. and Lincoln, Y. (1989). *Fourth Generation Evaluation*, Sage Publications: Newbury Park, California.

H

Hammersley, M. (2004). *Some questions about evidence-based practice,* in Thomas, G. and Pring, R. (Eds.) (2004) E*vidence-based practice in education*, Open University Press: New York.

Hargreaves, M. and Grenfell, A. (2003). 'The Use of Assessment Strategies to Develop Critical Thinking Skills in Science', paper presented to The Evaluation and Assessment Conference, Adelaide, South Australia, 24-25 November, 2003. [Accessed 21/10/07 from www.unisa.edu.au/evaluations/Full-papers/HargreavesFull.doc]

Hattie, J., Biggs, J. and Purdie, N. (1996). Effects of Learning Skills Interventions on Student Learning: A Meta-Analysis, *Review of Educational Research*, 66(2): 99-136 [Accessed 10/10/07 from Proquest 5000]

Hernandez, S. (2002). Team Learning in a Marketing Principles Course: Cooperative Structures That Facilitate Active Learning and Higher Level Thinking. *Journal of Marketing Education*, 24(1): 73-85 [Accessed 2/10/2004 from: http://online.sagepub.com].

House, E. (1993). *Professional evaluation: Social impact and political consequences*, in Mertens, D. (2005) *Research and Evaluation in Education and Psychology: Integrating Diversity with Quantitative, Qualitative, and Mixed Methods*, Sage Publications: Thousand Oaks, California.

I

Isaac, S. and Michael, W. (1990). *Handbook in Research and Evaluation*, EdITS Publishers: California.

Ivie, S. D. (1998). Ausubel's learning theory: An approach to teaching higher order thinking skills, *The High School Journal* 82(1): pp. 35-42.

J

James, R., McInnis, C. and Devlin, M. (2002). *Assessing Learning in Australian Universities*, Centre for the Study of Higher Education, University of Melbourne and The Australian Universities Teaching Committee: Melbourne. [Accessed 9/10/07 from www.cshe.inimelb.edu.au/assessinglearning]

Johnson, D., Johnson, R. and Smith, K. (1998). Cooperative Learning Returns To College: What Evidence Is There That It Works? *Change,* July/August, 1998: 27-35. [Accessed 18/07/07 from: http://academicadvising.studentservices.dal.ca/Files/Cooperative%20Learning.pdf]

Johnstone, C. (n.d.) *Fostering Deeper Learning*. [Accessed 10/04/2004 from: http://www.economics.inimelb.edu.au/TLdevelopment/fost.htm]

Joint Committee on Standards for Educational Evaluation (2007) *Draft Program Evaluation Standards Statement (Draft 3)*. [Accessed 28/11/07 from the Evaluation Center, Western Michigan University, at: http://www.wmich.edu/evalctr/jc/]

Jones, D. (1986) ‘Evaluation in adult education: some points for discussion’, in the Conference Proceedings of the Annual SCUTREA Conference 1986. [Accessed 8/7/07 from http://www.leeds.ac.uk/educol/documents/00002677.htm]

Jordan, J. (2007) The Application of Statistics Education Research in My Classroom, *Journal of Statistics Education*, 15(2). [Accessed 21/09/08 from http://www.amstat.org/publications/jse/v15n2/jordan.html]

Juwah, C. (2003). Using Peer Assessment to Develop Skills and Capabilities, *USDLA Journal* 17(1) [Accessed 11/01/2006 from: http://www.usdla.org/html/journal/JAN03_Issue/article04.html]

K

Keeton, M.and Tate, P. (1978). *Learning by Experience: what, why, how*, Jossey-Bass: SanFrancisco. Cited in Kolb, D. A. (1984). *Experiential Learning: experience as a source of learning and development,* Prentice-Hall: Englewood Cliffs.

Keller, J. (2007). *Motivation Design*. [Accessed 23/10/07 from http://www.arcsmodel.com/Mot%20dsgn%20Mot%20dsgn.htm]

Klenowski, V., Askew, S. and Carnell, E. (2006). Portfolios for learning, assessment and professional development in higher education, *Assessment and Evaluation in Higher Education* 31(3): 267-286.

Kolb, D. A. (1984). *Experiential Learning: experience as a source of learning and development,* Prentice-Hall: Englewood Cliffs.

Kuh, G. (2003). What we're learning about student engagement from NSSE. *Change*, 35(2), 24-33.

______(2003) The National Survey of Student Engagement: Conceptual Framework and Overview of Psychometric Properties. [Accessed 4/12/2003 from: http://www.iub.edu/~nsse/2003_annnual_report/pdf/NSSE_2003_Framework.pdf]

L

Lincoln, Y. and Guba, E. (2000). Paradigmic controversies, contradictions, and emerging confluences, as cited in Mertens, D. (2005) *Research and Evaluation in Education and Psychology: Integrating Diversity with Quantitative, Qualitative, and Mixed Methods*, (p.463) Sage Publications: Thousand Oaks, California.

Liu, E., Zhuo, Y.C. and Yuan, S.M. (2004). Assessing Higher-Order Thinking Using A Networked Portfolio System With Peer Assessment, *International Journal of Instructional Media* 31(2): 139-149.

Livingstone, D. and Lynch, K. (2000). Group Project Work and Student-Centred Active Learning: two different experiences, *Studies in Higher Education* 25(3): 325-345.

Lugenbehl, D. (2003). Learning at a Deeper Level, *Learning Abstracts* 6(2). [Accessed 10/04/05 from: http://www.league.org/publication/abstracts/learning/lelabs0302.htm]

M

Mainemelis, C., Boyatzis, R. and Kolb, D. (2002). Learning Styles and Adaptive Flexibility: Testing Experiential Learning Theory, *Management Learning*, 33(1): 5-33. [Accessed 4/11/04 from http://www.online.sagepub.com]

Marton, F. and Säljö, R (1976). On qualitative differences in learning. 1: Outcome and process, *British Journal of Educational Psychology,* 46: 4-11.

Masters, G. (2002) *Fair and meaningful measures? A review of examination procedures in the NSW Higher School Certificate*, Australian Council for Educational Research Ltd: Camberwell, Victoria. [Accessed online 23/10/07 from http://www.boardofstudies.nsw.edu.au/manuals/pdf_doc/masters_review.pdf]

McElwee, G. and Evans, A. (1992). Personal development portfolios: A mechanism for integrating quality in business education. *Education & Training* 34(7): 24-30.

McInerney, D. and McInerney, V. (1994). *Educational Psychology: constructing learning*, Prentice Hall: Sydney.

Melano, P. (2006). Measuring the student experience: Papers for the UOW Student Surveys Working Party, September 2006, University of Wollongong.

Mertens, D. (2005). *Research and Evaluation in Education and Psychology: Integrating Diversity with Quantitative, Qualitative, and Mixed Methods*, Sage Publications: Thousand Oaks, California.

Mohammed, S., Klimoski, R. and Rentsch, J. R. (2000). The Measurement of Team Mental Models: We Have No Shared Schema, *Organizational Research Methods* 3(2): 123-165. [Accessed 12/10/2004 from: http://online.sagepub.com]

Montgomery, K. (2002). Authentic Tasks and Rubrics: Going Beyond Traditional Assessments in College Teaching, *College Teaching*, 50(1): 34-39. [Accessed through ProQuest 2/7/03]

Moore, D. and Cobb, G. (2000). Statistics and Mathematics: Tension and Cooperation, *The American Mathematical Monthly*, 107(7): 615-630. [Accessed from ProQuest 23/1/06]

Morris, M., Porter, A. and Griffiths, D. (2007a). 'Signposting Learning: using defined learning outcomes to facilitate alignment of teaching, learning and assessment', in *Proceedings of REAP International Online* Conference, May, 2007 [Available at: http://www.reap.ac.uk].

______(2007b). *Using the end to justify the means: enabling student learning of statistics*. International Journal of Learning Volume 13(10) 157-168 [Available at: http://www.Learning_Journal.com].

______(2006a). 'Using the end to justify the means: enabling student learning of statistics'. *Presented at the 13th International Conference on Learning*, Montego Bay, Jamaica, June, 2006.

______ (2006b). Trawling for deeper learning: achieving an optimum catch, *International Journal of Learning*, Volume 13 (5), pp.137-146.

______(2005a). 'Communicating Understanding in Statistics: An Exercise in Deeper Learning of Statistical Concepts'. Presented at the 2005 International Statistics Institute Conference, Darling Harbour, Sydney, April, 2005.

______(2005b). 'Assessment: the end or a good place to start'. Presented at The Sixth Annual J. B. Douglas Postgraduate Awards Day, Macquarie Graduate School of Management, November, 2005.

______ (2004a). Assessment is Bloomin' Luverly: Developing Assessment that Enhances Learning. *Journal of University Teaching and Learning Practice*, 1(2), 90-106. [Available from: http://jutlp@uow.edu.au]

______(2004b) *Getting Down and Dirty in Statistics*, Proceedings of the *2004 Workshop on Research Methods: Statistics and Finance*. University of Wollongong School of Mathematics and Applied Statistics: pp. 173-186.

______(2004c). 'Assessment is Bloomin' Luverly'. Presented at ARTIST Roundtable Conference, Appleton, Wisconsin, August, 2004.

______(2003). 'Assessment as a Tool for Learning'. Presented at the Evaluations and Assessment Conference, Adelaide, S.A., November, 2003.

N

National Committee of Inquiry into Higher Education (NCIHE, also known as the Dearing Committee) (1997) *Higher Education in the Learning Society*, Middlesex: Hayes. [accessed through Google on 16/10/07 at: http://www.leeds.ac.uk/educol/ncihe/]

Nightingale, P., Te Wiata, I., Toohey, S. Ryan, G., Hughes, C. and Magin, D. (1996) *Assessing learning in Universities*, University of New South Wales Press: Kensington, New South Wales.

Nieweg, M. R. (2000). Learning to Reflect: a practical theory of teaching, Amsterdam University of Professional Education: Netherlands.
Available from: http://www.oro.hva.nl/files/attachments/1p_refl-nieweg.306-0.pdf [accessed:26/06/04]

Nuhfer, E. and Knipp, D. (2003) The Knowledge Survey: A Tool for All Reasons, *To Improve the Academy*, 21: 59-78. [Accessed 21/09/08 from http://www.isu.edu/ctl/facultydev/KnowS.htm]

O

O'Leary, Z. (2005). *The Essential Guide to Doing Research*, Sage Publications: London.

Osborn, M. and Plunkett, M. (2003). 'Assessment and learning styles; mix and match or a mismatch?' Presented to the NZARE/AARE Joint Conference 29 November-3 December, 2003, Auckland, New Zealand.

Oxford Advanced Learner's Dictionary (2005) accessed 31/08/07 from http://www.oup.com/oald-bin/web_getald7index1a.pl.

P

Patton, M. (2002) *Qualitative Research & Evaluation Methods*, in Mertens, D. (2005) *Research and Evaluation in Education and Psychology: Integrating Diversity with Quantitative, Qualitative, and Mixed Methods*, Sage Publications: Thousand Oaks, California.

Pellegrino, J. W., Baxter, G. F. and Glaser, R. (1999) Addressing the "Two Disciplines" problem: linking theories of cognition with assessment and instructional, practice, *Review of Research in Education* 24(1): 307-353 [Accessed 21/10/07 from http://rre.sagepub.com.ezproxy.uow.edu.au:2048/cgi/reprint/24/1/307]

Pfaff, E. and Huddlestone, P. (2003). Does It Matter If I Hate Teamwork? What Impacts Student Attitudes toward Teamwork, *Journal of MarketingEducation* 25(1): 37-45. [Accessed 12/10/2004 from: http://online.sagepub.com]

Pintrich, P. and De Groot, E. (1990). Motivational and Self-Regulated Learning Components of Classroom Academic Performance, *Journal of Educational Psychology*, 82(1): p.33-40.

Piper, D., Nulty, D. and O'Grady, G. (1996). *Examination Practices and Procedures in Australian Universities,* Higher Education Division DEETYA Evaluations and Investigations Program, Australian Government Publishing Service (AGPS): Canberra.

Punch, M. (1986). *The Politics and Ethics of Fieldwork*, in Fontana, A. and Frey, J. (1994). *Interviewing: The Art of Science*, in Denzin, N. and Lincoln, Y. *Handbook of qualitative research*, (pp. 361-376). Sage Publications: Newbury Park, California.

R

Ramsden, P. (1992). *Learning to Teach in Higher Education*, Routledge: New York.

Riding, R. and Rayner, S. (1999). *Learning Enhancement Programme – Cognitive Style and Effective Teaching,* Learning & Training Technology: Birmingham.

Robson, C. (2002). *Real World Research: A Resource for Social Scientists and Practitioner- Researchers,* 2nd Edition. Blackwell Publishing: Carlton, Australia.

Rosenshine, B. and Meister, C. (1992). The Use of Scaffolds for Teaching Higher-Level Cognitive Strategies, *Educational Leadership,* 49(7): 26-33.

Rossman, A. (1997). Quantitative Reasoning: Argument with Data, *College Teaching*, 45(2): 52-54. [Accessed from ProQuest 23/1/06]

Rossman, A. and Chance, B. (1999). Teaching the reasoning of statistical inference: A "top ten" list, *The College Mathematics Journal*, 30(4): 297-306.

Rumsey, D. (2002). Statistical Literacy as a Goal for Introductory Statistics Courses, *Journal of Statistics Education*, 10(3). [Accessed 3/7/2008 from http://amstat.org.ezproxy.uow.edu:2048/publications/jse/v10n3/rumsey.html]

S

Sadler, R. (1998). Formative Assessment: revisiting the territory, *Assessment in Education* 5(1): 77-84 [Accessed 9/10/07 from Proquest]

Sebatane, E. (1998). Assessment and Classroom Learning: a response to Black and Wiliam, *Assessment in Education* 5(1): 123-130 [Accessed 9/10/07 from Proquest]

Scriven, M. (2005). 'The Logic and Methodology of Checklists'. [Accessed 28/11/07 from the Evaluation Centre, University of Western Michigan at: http://www.wmich.edu/evalctr/jc/]

Simonson, M. and Abu-Omar, K. (eds.) (1993). *Annual Proceedings of selected research and development presentations at the 1993 national convention of the Association for Educational Communications and Technology*, New Orleans, pp.1047-1080.

Stake R. (1995). *The Art of Case Study* Research, Sage Publications: Thousand Oaks, California.

Statistical Society of Australia Inc. (SSAI) (2005). *The State of Statistics in Australian Universities* [Accessed 24/1/08 from http://www.statsoc.org.au/review0405/ReviewofStatsFinalReport.pdf]

______(2004). The State of Statistics in Australian Universities – an SSAI-sponsored Review, *SSAI Newsletter*, October, 2004, pp.7-9.

Stufflebeam D. (2005). Guidelines for Developing Evaluation Checklists: The Checklists development Checklist (CDC). [Accessed 28/11/07 from the Evaluation Centre, University of Western Michigan at: http://www.wmich.edu/evalctr/jc/]

T

Tashakkori, A. and Teddlie, C. (1998). *Mixed Methodology*, in Mertens, D. (2005) *Research and Evaluation in Education and Psychology: Integrating Diversity with Quantitative, Qualitative, and Mixed Methods*, Sage Publications: Thousand Oaks, California.

Te Wiata, I. (1996). *Thinking critically and making judgements*, in Nightingale, P., Te Wiata, I., Toohey, S. Ryan, G., Hughes, C. and Magin, D., *Assessing learning in Universities*, University of New South Wales Press: Kensington, New South Wales, pp.14-38.

Thomas, G. and Pring, R. (Eds.) (2004). E*vidence-based practice in education*, Open University Press: New York.

Toohey, S. (1996). *Thinking critically and making judgements*, in Nightingale, P., Te Wiata, I., Toohey, S. Ryan, G., Hughes, C. and Magin, D., *Assessing learning in Universities*, University of New South Wales Press: Kensington, New South Wales, pp.79-114.

Trigwell, K., Prosser, M. and Waterhouse, F. (1999). Relations between teachers' approaches to teaching and students' approaches to learning, *Higher Education* 35: 57-70.

Tucker, S. (1993). *Evaluating interactive technologies: A cognitive model, in Simonson*, M. and Abu-Omar, K. (eds.) (1993) *Annual Proceedings of selected research and development presentations at the 1993 national convention of the Association for Educational Communications and Technology*, New Orleans, pp.1047-1080.

U

University of Western Sydney, n.d., 'Facts and figures'. [Accessed 5/05/07 from http://policies.uws.edu.au]

University of Wollongong (2008). 'UOW Graduate Qualities', [accessed 20/06/08 from: http://www.uow.edu.au/about/teaching/qualities/]

______ n.d.(a), 'Facts and figures'. [Accessed 8/05/08 from http://www.uow.edu.au/about/facts/index.html]

______ n.d.(b), ‘Attributes of a UOW Graduate’. [accessed 14/05/08 from http://www.uow.edu.au/about/teaching/graduate_attributes.html]

Ussher, B. and Gibbes, C. (2002). Vygotsky, Physical Education and Social Interaction, *Journal of Physical Education New Zealand*, July 35:1 76-86.

W

Watkins, C., Carnell, E., Lodge, C., Wagner, P. and Whalley, C. (2002). Effective Learning, *National School Improvement Network Research Matters*, No. 13. Institute of Education: London. Available from: http://www.ioe.ac.uk/iseic/research.pdf [Accessed 4/7/07]

Wehmeier, S. (Ed.) (2005). *Oxford Advanced Learner's Dictionary*, Oxford University Press, 2005 [accessed 29/08/07 from http://www.oup.com/oald-bin/web_getald7index1a.pl]

West, R. (1998). *Learning for Life: Review of Higher Education Financing and Policy*, sponsored by the department of Employment, Training and Youth Affairs, Commonwealth of Australia. [Accessed 16/08/07 from http://www.dest.gov.au/archive/highered/hereview/]

Wild, C. and Pfannkuch, M. (1999). Statistical Thinking in Empirical Enquiry, *International Statistics Review*, 67(3): 223-248.

Wirth K., and Perkins D. (2006). *Knowledge Surveys: An Indispensable Course Design and Assessment Tool*. [Presented at *Innovations in the Scholarship of Teaching and Learning at Liberal Arts Colleges* and accessed 21/09/08 from http://www.macalester.edu/geology/WirthPerkinsKS.pdf]

Wood, D. and Wood, H. (1996). Vygotsky, Tutoring and Learning, *Oxford Review of Education*, 22(1): 5-16. [Accessed from Academic Search Premier 5/3/07]

Y

Yorke, M. (2001) Formative Assessment and its Relevance to Retention, *Higher Education Research & Development,* 20(2): 115-126.

Appendices

Appendix 3.1a: Table of classifications of 'lower order' knowledge types and cognitive processing skills according to the revised Bloom's Taxonomy[1]

Lower order knowledge types[2]	Lower order cognitive processing skills[2]
1. Facts	1. Remember
1.1 Terminology	1.1 Identify
1.2 Details	1.2 Recall
2. Concepts	2. Understand
2.1 Classifications	2.1 Interpret
3. Procedures	2.2 Exemplify
3.1 Subject specific skills	2.3 Classify
3.2 Subject specific techniques	3. Apply
	3.1 Execute
	3.2 Implement

[1] Anderson and Krathwohl, 2001.

[2] These knowledge types and cognitive skills have been identified by the teacher as 'lower/higher order'. The 'lower order' knowledge and skills assume greater focus in high school teaching environments than in universities while the reverse should be the focus at universities.

Appendix 3.1b: Table of classifications of 'higher order' knowledge types and cognitive processing skills according to the revised Bloom's Taxonomy[1]

Higher order knowledge types[2]	Higher order cognitive processings[2]
2. Concepts	2. Understand
2.2 Generalisations and principles	2.4 Summarise
2.3 Models	2.5 Infer
3. Procedures	2.6 Compare
3.3 Criteria for selecting procedures	2.7 Explain
4. Metacognition	4. Analyse
4.1 Strategies	4.1 Differentiate
4.2 Context	4.2 Organise
4.3 Self-knowledge	4.3 Attribute
	5. Evaluate
	5.1 Check
	5.2 Critique
	6. Create
	6.1 Generate
	6.2 Plan
	6.3 Produce

[1] Anderson and Krathwohl, 2001.

[2] These knowledge types and cognitive skills have been identified by the teacher as 'lower/higher order'. The 'lower order' knowledge and skills assume greater focus in high school teaching environments than in universities while the reverse should be the focus at universities.

Appendix 3.2: Model of Statistical Reasoning

Please see print copy for Appendix 3.2

Source: Garfield, 2002, p. 8

Appendix 5.1: The Metfessel-Michael Paradigm

Please see print copy for Appendix 5.1

Source: Isaac and Michael, 1990, pp. 15-16

Appendix 5.1: The Metfessel-Michael Paradigm (continued)

Please see print copy for Appendix 5.1

Source: Isaac and Michael, 1990, pp. 17-18

Appendix 5.2: The Four Step Evaluation Model

Please see print copy for Appendix 5.2

Source: Isaac and Michael, 1990, pp. 17-18

Appendix 5.2: The Four Step Evaluation Model (continued)

Please see print copy for Appendix 5.2

Source: Isaac and Michael, 1990, pp. 17-18

Appendix 5.3: An integrated evaluation framework (adapted from Alexander & Hedberg, 1994)

Please see print copy for Appendix 5.3

Source: Bain, 1999, pp. 168-169

Appendix 5.3: An integrated evaluation framework (adapted from Alexander & Hedberg, 1994) (continued)

Please see print copy for Appendix 5.3

Source: Bain, 1999, pp. 168-169

Appendix 6.1

STAT 131: Aspects of ***lectures***[1] and associated student survey responses across all implementations

Aspect description	Autumn, 2003	Autumn, 2004	Spring, 2004	Autumn, 2005	Spring, 2005
Coordinator/lecturer the subject designer	✓	✓	×	✓	×
Primarily expository	✓	Participation encouraged	✓	Participation encouraged	✓
Content aligned with lab tasks	✓	Content and thinking aligned with lab tasks	✓	Content and thinking aligned with lab tasks	✓
Lectures used both auditory and visual stimuli to involve more students	✓	✓	Lectures used PowerPoint and videoed material only	✓	Lectures used PowerPoint and videoed material only
Teacher/student interaction	✓	✓	Little interaction	✓	Little interaction
Non compulsory attendance but strongly recommended	✓	✓	✓	✓	✓
Total enrolment[2]	159	205	172	179	160
Number of survey responses[3]	63	101	63	62	21
Percentage of responses reporting attendance 10-13 weeks[3]	44	48	19	44	14
Percentage of responses regarding ***lectures*** as moderately to extremely important to learning[3]	76	73	56	79	57
Ranking of importance to student learning[4]	5	8	11	7	12

[1] Source: Annotated journal
[2] Source: Student assessment data
[3] Source: Student survey data
The number of facets surveyed differed across the sessions: Autumn 2003 – 6 facets
Autumn 2004 – 14 facets
Spring 2004 – 16 facets
Autumn 2005 – 16 facets
Spring 2005 – 16 facets
[4] Data selected from Morris, Porter, and Griffiths, 2007b, p. 161

Appendix 6.2

STAT131: Aspects of ***lab classes***[1] and student survey responses across all implementations

Aspect description	Autumn, 2003	Autumn, 2004	Spring, 2004	Autumn, 2005	Spring, 2005
Class size: 25-30 students	✓	25-40 students (wide variations)	25-40 students (wide variations)	25-40 students (wide variations)	15-40 students (wide variations)
Non-compulsory attendance but completion of tasks required	✓	✓	✓	✓	✓
Collaborative learning encouraged	✓	Collaborative learning *actively promoted*	Collaborative learning *actively promoted*	Collaborative learning *actively promoted*	Collaborative learning *actively promoted*
Authentic tasks	✓	✓	✓	✓	✓
Tasks completed primarily in lab class and solutions to tasks maintained by students	✓	Tasks completed designed promote completion outside of lab class and solutions constructed within a lab manual	Tasks reduced in number	Task numbers reduced still further	Tasks reorganised
Solutions to tasks maintained by students	✓	Solutions constructed within a *lab manual*	Solutions constructed within a *lab manual*	Solutions constructed within a *lab manual*	Solutions constructed within a *lab manual*
No solutions to tasks made available	✓	Delayed release of task solutions to the web	Solution release was *irregular and too late to provide feedback before assignments*	Timing of release of solutions improved to meet assignment deadlines	Solution release was again *irregular and too late to provide feedback before assignments*
Tutor's role weakly defined: administrative/supportive	✓	Tutor's role more defined – pro-activity prescribed	Tutor's role more defined – pro-activity prescribed	Tutor's role more defined – pro-activity prescribed	Tutor's role more defined – pro-activity prescribed
Participation marks based on attendance and completion of tasks but not defined prescriptively	✓	Participation marks (assessment) based on completion of tasks but not defined prescriptively	Participation marks (assessment) based on completion of tasks but not defined prescriptively	Participation marks(assessment) based on completion of tasks but not defined prescriptively	Participation marks (assessment) based on completion of tasks but not defined prescriptively
Total enrolment[2]	159	205	172	179	160
Number of survey responses[3]	63	101	63	62	21
Percentage within implementation regarding ***lab classes*** as moderately to extremely important to learning[3]	91	87	76	82	86

Appendix 6.2 continued

STAT131: Aspects of ***lab classes*** and student survey responses across all implementations

Aspect description	Autumn, 2003	Autumn, 2004	Spring, 2004	Autumn, 2005	Spring, 2005
Ranking of importance to student learning[4]	2	4	8	5	6.5

[1] Source: Annotated journal
[2] Source: Student assessment data
[3] Source: Student survey data
The number of facets surveyed differed across the sessions: Autumn 2003 – 6 facets
Autumn 2004 – 14 facets
Spring 2004 – 16 facets
Autumn 2005 – 16 facets
Spring 2005 – 16 facets
[4] Data selected from Morris, Porter, and Griffiths, 2007b, p. 161

Appendix 6.3

STAT131: Aspects of ***lab manuals***[1] and student survey responses across all implementations

Aspect description	Autumn, 2003	Autumn, 2004	Spring, 2004	Autumn, 2005	Spring, 2005
Manual contains tasks, practice questions and lecture notes only	✓	Manual contains tasks, *past exam papers, SPSS notes, selected statistical tables* and lecture notes	Manual contains tasks, *past exam papers, SPSS notes, selected statistical tables* and lecture notes	Manual contains tasks, *past exam papers, SPSS notes, selected statistical tables* and lecture notes	Manual contains tasks, *past exam papers, SPSS notes, selected statistical tables* and lecture notes
No provision for student responses	×	Provision for student responses	Provision for student responses	Provision for student responses	Provision for student responses
Manual provided *portfolio* of learning which students took into the final exam	×	✓	✓	✓	✓
Learning framework exercises included in the tasks	×	×	×		
Subject unit outline included in the manual introduction	×	✓	✓	✓	✓
Assignments included in the manual	×	✓	×	First assignment only included	×
Expected time to complete tasks displayed for each week	×	✓	✓	✓	✓
Warnings of impending assessment displayed one week in advance	×	×	✓	✓	✓
Total enrolment[2]	159	205	172	179	160
Number of survey responses[3]	63	101	63	62	21

Appendix 6.3 continued
STAT131: Aspects of ***lab manuals*** and student survey responses across all implementations

Aspect description	Autumn, 2003	Autumn, 2004	Spring, 2004	Autumn, 2005	Spring, 2005
Percentage within implementation regarding ***lab manuals*** as moderately to extremely important to learning[3]	No lab manual	92	89	92	95
Ranking of importance to student learning[4]	-	2	4	2	3.5

[1] Source: Annotated journal
[2] Source: Student assessment data
[3] Source: Student survey data
The number of facets surveyed differed across the sessions: Autumn 2003 – 6 facets
Autumn 2004 – 14 facets
Spring 2004 – 16 facets
Autumn 2005 – 16 facets
Spring 2005 – 16 facets
[4] Data selected from Morris, Porter, and Griffiths, 2007b*,* p. 161

Appendix 6.4
STAT131: Aspects of ***lab tasks***[1] and student survey responses across all implementations

Aspect description	Autumn, 2003	Autumn, 2004	Spring, 2004	Autumn, 2005	Spring, 2005
Task learning objectives defined	✗	✓	✓	✓	✓
Data files constructed from class-based student measurement	✗	✓	✓	✓	✓
Survey data collected on the web	✓	✓	✓	✓	✓
SPSS based tasks	✓	✓	✓	✓	✓
Pen and paper exercises	Included as supplementary tasks in the lab manual	Interspersed with practical tasks	Interspersed with practical tasks	Interspersed with practical tasks	Interspersed with practical tasks
Task solutions	Not generally available	Delayed release to web	Release was inadequate for feedback before assignment due dates	Some issues with delayed release of some sets of solutions	Release was inadequate for feedback before assignment due dates
Tasks completed outside class time	✗	✓	✓	✓	✓
Total enrolment[2]	159	205	172	179	160
Number of survey responses[3]	63	101	63	62	21
Percentage within implementation regarding ***lab tasks*** as moderately to extremely important to learning[3]	79	87	89	89	97
Ranking of importance to student learning[4]	4	5	3	3.5	2

[1] Source: Annotated journal
[2] Source: Student assessment data
[3] Source: Student survey data
The number of facets surveyed differed across the sessions: Autumn 2003 – 6 facets
Autumn 2004 – 14 facets
Spring 2004, autumn 2005 and spring 2005 – 16 facets
[4] Data selected from Morris, Porter, and Griffiths, 2007b, p. 161

Appendix 6.5
STAT131: Aspects of ***solutions***[1] and student survey responses across all implementations

Aspect Description	Autumn, 2003	Autumn, 2004	Spring, 2004	Autumn, 2005	Spring, 2005
Solutions to lab tasks Solutions to assessment	Answers only in lab manual Released in hard copy one week after submission date No detailed marking criteria provided	Delayed availability online but released before associated assessment Released online week after submission date Detailed marking criteria provided	Irregularly released online, sometimes too late for associated assessment Released online week after submission date Detailed marking criteria provided	Delayed availability online but released before associated assessment Released online week after submission date Detailed marking criteria provided	Delayed availability online but released before associated assessment Released online week after submission date Detailed marking criteria provided
Total enrolment[2]	159	205	172	179	160
Number of survey responses[3]	63	101	63	62	21
Percentage within implementation regarding ***solutions*** as moderately to extremely important to learning[3]	Not surveyed	90	94	97	100
Ranking of importance to student learning[4]	-	3	2	1	1

[1] Source: Annotated journal
[2] Source: Student assessment data
[3] Source: Student survey data
The number of facets surveyed differed across the sessions: Autumn 2003 – 6 facets
Autumn 2004 – 14 facets
Spring 2004 – 16 facets
Autumn 2005 – 16 facets
Spring 2005 – 16 facets
[4] Data selected from Morris, Porter, and Griffiths, 2007b*,* p. 161

Appendix 6.6

STAT131: Aspects of ***assignments***[1] and student survey responses across all implementations

Aspect Description	Autumn, 2003	Autumn, 2004	Spring, 2004	Autumn, 2005	Spring, 2005
Teamwork (pairs)	Some common questions	No common questions	No common questions	No common questions	No common questions
Parallel pairs of questions (same question different data sets/models)	✓	✓	✓	✓	✓
Complementary questions (one students calculations based upon the partner's answers)	✗	✓	✓	✓	✓
Total enrolment[2]	159	205	172	179	160
Number of survey responses[3]	63	101	63	62	21
Percentage within implementation regarding ***assignments*** as moderately to extremely important to learning[3]	94	93	95	89	95
Ranking of importance to student learning[4]	1	1	1	3.5	3.5

[1] Source: Annotated journal
[2] Source: Student assessment data
[3] Source: Student survey data
The number of facets surveyed differed across the sessions: Autumn 2003 – 6 facets
Autumn 2004 – 14 facets
Spring 2004 – 16 facets
Autumn 2005 – 16 facets
Spring 2005 – 16 facets
[4] Data selected from Morris, Porter, and Griffiths, 2007b, p. 161

Appendix 6.7 STAT131: Aspects of ***lecture notes***[1] and student survey responses across all implementations					
Aspect Description	Autumn, 2003	Autumn, 2004	Spring, 2004	Autumn, 2005	Spring, 2005
Lecture notes available in lab manual	✓	✓	✓	✓	✓
Lecture notes available online	✓	✓	✓	✓	✓
Total enrolment[2]	159	205	172	179	160
Number of survey responses[3]	63	101	63	62	21
Percentage within implementation regarding ***lecture notes*** as moderately to extremely important to learning[3]	Not surveyed	86	87	79	86
Ranking of importance to student learning[4]	-	6	5	7	6.5

[1] Source: Annotated journal
[2] Source: Student assessment data
[3] Source: Student survey data
The number of facets surveyed differed across the sessions: Autumn 2003 – 6 facets
Autumn 2004 – 14 facets
Spring 2004 – 16 facets
Autumn 2005 – 16 facets
Spring 2005 – 16 facets
[4] Data selected from Morris, Porter, and Griffiths, 2007b*,* p. 161

Appendix 6.8

STAT131: Aspects of ***marking guides***[1] and student survey responses across all implementations

Aspect Description	Autumn, 2003	Autumn, 2004	Spring, 2004	Autumn, 2005	Spring, 2005
Available with assignments	✗	✓	✓	✓	✓
Order aligned with assignment questions	✗	✓	✓	✓	✓
Marks recorded as 0/1	✗	✗	✓	✓	✓
Criteria aligned with lecture notes and lab tasks	✗	✗	✓	✓	✓
Criteria consistent internally	n/a	✗	✗	✓	✓
Total enrolment[2]	159	205	172	179	160
Number of survey responses[3]	63	101	63	62	21
Percentage within implementation regarding ***lecture notes*** as moderately to extremely important to learning[3]	Not surveyed	65	70	66	76
Ranking of importance to student learning[4]	-	10	9	12	9

[1] Source: Annotated journal
[2] Source: Student assessment data
[3] Source: Student survey data
The number of facets surveyed differed across the sessions: Autumn 2003 – 6 facets
Autumn 2004 – 14 facets
Spring 2004 – 16 facets
Autumn 2005 – 16 facets
Spring 2005 – 16 facets
[4] Data selected from Morris, Porter, and Griffiths, 2007b*,* p. 161

Appendix 6.9
STAT131: Aspects of ***midterm***[1] and student survey responses across all implementations

Aspect Description	Autumn, 2003	Autumn, 2004	Spring, 2004	Autumn, 2005	Spring, 2005
Students prepared summary notes that were marked	✓	✗	✗	✗	✗
Questions aligned with lecture notes	✓	✓	✓	✓	✓
Questions aligned with lab tasks	✗	✗	✗	✓	✓
Total enrolment[2]	159	205	172	179	160
Number of survey responses[3]	63	101	63	62	21
Percentage within implementation regarding ***midterm*** as moderately to extremely important to learning[3]	83	74	81	71	57
Ranking of importance to student learning[4]	3	7	7	10	13

[1] Source: Annotated journal
[2] Source: Student assessment data
[3] Source: Student survey data
The number of facets surveyed differed across the sessions: Autumn 2003 – 6 facets
Autumn 2004 – 14 facets
Spring 2004 – 16 facets
Autumn 2005 – 16 facets
Spring 2005 – 16 facets
[4] Data selected from Morris, Porter, and Griffiths, 2007b, p. 161

Appendix 6.10
STAT131: Aspects of ***online lecture notes***[1] and student survey responses across all implementations

Aspect Description	Autumn, 2003	Autumn, 2004	Spring, 2004	Autumn, 2005	Spring, 2005
Lecture notes available online	✓	✓	✓	✓	✓
Total enrolment[2]	159	205	172	179	160
Number of survey responses[3]	63	101	63	62	21
Percentage within implementation regarding ***online lecture notes*** as moderately to extremely important to learning[3]	46	54	84	79	91
Ranking of importance to student learning[4]	6	11	6	7	5

[1] Source: Annotated journal
[2] Source: Student assessment data
[3] Source: Student survey data
The number of facets surveyed differed across the sessions: Autumn 2003 – 6 facets
Autumn 2004 – 14 facets
Spring 2004 – 16 facets
Autumn 2005 – 16 facets
Spring 2005 – 16 facets
[4] Data selected from Morris, Porter, and Griffiths, 2007b, p. 161

Appendix 6.11
STAT131: Aspects of ***teamwork***[1] and student survey responses across all implementations

Aspect Description	Autumn, 2003	Autumn, 2004	Spring, 2004	Autumn, 2005	Spring, 2005
Teamwork a focus in lab classes	×	✓	✓	✓	✓
Assignment teamwork using common questions	✓	✓	×	×	×
Assignment teamwork using parallel questions	✓	✓	✓	✓	✓
Assignment teamwork using complementary questions	×	✓	✓	✓	✓
Total enrolment[2]	159	205	172	179	160
Number of survey responses[3]	63	101	63	62	21
Percentage within implementation regarding ***teamwork*** as moderately to extremely important to learning[3]	Not surveyed	Not surveyed	65	68	72
Ranking of importance to student learning[4]	-	-	10	11	10

[1] Source: Annotated journal
[2] Source: Student assessment data
[3] Source: Student survey data
The number of facets surveyed differed across the sessions: Autumn 2003 – 6 facets
Autumn 2004 – 14 facets
Spring 2004 – 16 facets
Autumn 2005 – 16 facets
Spring 2005 – 16 facets
[4] Data selected from Morris, Porter, and Griffiths, 2007b, p. 161

Appendix 6.12
STAT131: Aspects of ***tutor***[1] and student survey responses across all implementations

Aspect Description	Autumn, 2003	Autumn, 2004	Spring, 2004	Autumn, 2005	Spring, 2005
Tutor role generally perceived as passive facilitator	✓	✓	×	×	✓
Tutor engagement in teaching actively encouraged	×	✓	✓	✓	✓
Total enrolment[2]	159	205	172	179	160
Number of survey responses[3]	63	101	63	62	21
Percentage within implementation regarding ***tutor*** as moderately to extremely important to learning[3]	Not surveyed	71	21	77	81
Ranking of importance to student learning[4]	-	9	16	9	8

[1] Source: Annotated journal
[2] Source: Student assessment data
[3] Source: Student survey data
The number of facets surveyed differed across the sessions: Autumn 2003 – 6 facets
Autumn 2004 – 14 facets
Spring 2004 – 16 facets
Autumn 2005 – 16 facets
Spring 2005 – 16 facets
[4] Data selected from Morris, Porter, and Griffiths, 2007b, p. 161

Appendix 6.13
STAT131: Aspects of ***learning strategies***[1] and student survey responses across all implementations

Aspect Description	Autumn, 2003	Autumn, 2004	Spring, 2004	Autumn, 2005	Spring, 2005
Learning strategies specifically addressed in lectures	✗	✗	✓	✓	✓
Learning strategies included among lab tasks	✗	✗	✓	✓	✓
Learning strategies included as summary tasks in lab manual	✗	✗	✗	✗	✓
Total enrolment[2]	159	205	172	179	160
Number of survey responses[3]	63	101	63	62	21
Percentage within implementation regarding ***learning strategies*** as moderately to extremely important to learning[3]	Not surveyed	Not surveyed	35	65	67
Ranking of importance to student learning[4]	-	-	15	13	11

[1] Source: Annotated journal
[2] Source: Student assessment data
[3] Source: Student survey data
The number of facets surveyed differed across the sessions: Autumn 2003 – 6 facets
Autumn 2004 – 14 facets
Spring 2004 – 16 facets
Autumn 2005 – 16 facets
Spring 2005 – 16 facets
[4] Data selected from Morris, Porter, and Griffiths, 2007b*,* p. 161

Appendix 6.14
STAT131: Aspects of ***objectives***[1] and student survey responses across all implementations

Aspect Description	Autumn, 2003	Autumn, 2004	Spring, 2004	Autumn, 2005	Spring, 2005
Objectives specified for lectures	✓	✓	✓	✓	✓
Objectives specified for lab tasks	✕	✓	✓	✓	✓
Objectives specified for assignments	✕	✓	✓	✓	✓
Total enrolment[2]	159	205	172	179	160
Number of survey responses[3]	63	101	63	62	21
Percentage within implementation regarding ***objectives*** as moderately to extremely important to learning[3]	Not surveyed	44	48	50	48
Ranking of importance to student learning[4]	-	13	12	14	14

[1] Source: Annotated journal
[2] Source: Student assessment data
[3] Source: Student survey data
The number of facets surveyed differed across the sessions: Autumn 2003 – 6 facets
Autumn 2004 – 14 facets
Spring 2004 – 16 facets
Autumn 2005 – 16 facets
Spring 2005 – 16 facets
[4] Data selected from Morris, Porter, and Griffiths, 2007b, p. 161

Appendix 6.15
STAT131: Aspects of ***forum***[1] and student survey responses across all implementations

Aspect Description	Autumn, 2003	Autumn, 2004	Spring, 2004	Autumn, 2005	Spring, 2005
Online student discussion of subject related issues	✓	✓	✓	✓	✓
Total enrolment[2]	159	205	172	179	160
Number of survey responses[3]	63	101	63	62	21
Percentage within implementation regarding ***forum*** as moderately to extremely important to learning[3]	Not surveyed	52	37	45	14
Ranking of importance to student learning[4]	-	12	13.5	15	16

[1] Source: Annotated journal
[2] Source: Student assessment data
[3] Source: Student survey data
The number of facets surveyed differed across the sessions: Autumn 2003 – 6 facets
Autumn 2004 – 14 facets
Spring 2004 – 16 facets
Autumn 2005 – 16 facets
Spring 2005 – 16 facets
[4] Data selected from Morris, Porter, and Griffiths, 2007b*,* p. 161

Appendix 6.16
STAT131: Aspects of ***text***[1] and student survey responses across all implementations

Aspect Description	Autumn, 2003	Autumn, 2004	Spring, 2004	Autumn, 2005	Spring, 2005
Text prescribed for all students	✓	✓	✓		
Text listed as a reference only	×	×	×	×	✓
Total enrolment[2]	159	205	172	179	160
Number of survey responses[3]	63	101	63	62	21
Percentage within implementation regarding ***text*** as moderately to extremely important to learning[3]	Not surveyed	28	37	29	24
Ranking of importance to student learning[4]	-	14	13.5	16	15

[1] Source: Annotated journal
[2] Source: Student assessment data
[3] Source: Student survey data
The number of facets surveyed differed across the sessions: Autumn 2003 – 6 facets
Autumn 2004 – 14 facets
Spring 2004 – 16 facets
Autumn 2005 – 16 facets
Spring 2005 – 16 facets
[4] Data selected from Morris, Porter, and Griffiths, 2007b*,* p. 161

The following notes apply to Appendices 6.1 – 6.16

[1] Source: Annotated journal
[2] Source: Student assessment data
[3] Source: Student survey data

The number of facets surveyed differed across the sessions: Autumn 2003 – 6 facets
Autumn 2004 – 14 facets
Spring 2004 – 16 facets
Autumn 2005 – 16 facets
Spring 2005 – 16 facets

[4] Data selected from Morris, Porter, and Griffiths, 2007b, p. 161

Appendix 6.17
STAT131: Facets of the teaching/learning framework and student survey responses to perceived importance to learning across all implementations

Facet	Autumn, 2003		Autumn, 2004		Spring, 2004		Autumn, 2005		Spring, 2005	
	Percentage[1]	Ranking[2]	Percentage[1]	Ranking[2]	Percentage[1]	Ranking[2]	Percentage[1]	Ranking[2]	Percentage[1]	Ranking[2]
Solutions	[3]	-	90	3	94	2	97	1	100	1
Assignments	94	1	93	1	95	1	89	3.5	95	3.5
Lab Manual	[3]	-	92	2	89	4	92	2	95	3.5
Lab tasks	79	4	87	5	89	3	89	3.5	89	2
Lecture Notes	[3]	-	86	6	87	5	79	7	86	6.5
Lab classes	90	3	87	4	76	8	82	5	86	6.5
Marking guides	[3]	-	65[4]	10	70[4]	9	66[5]	12	76[4]	9
Midterm quiz	[3]	2	83	7	74	7	81	10	57	13
Online lecture notes	46	6	54	11	84	6	79	7	91	5
Lectures	76	5	73	8	56	11	79	7	57	12
Teamwork	[3]	-	[3]	-	65	10	68	11	71	10
Tutor	[3]	-	71	9	21	16	77	9	81	8
Learning strategies	[3]	-	[3]	-	35	15	65	13	67	11
Objectives	[3]	-	44	13	48	12	50	14	47	14
Forum	[3]	-	52	12	37	13.5	45	15	15	16
Textbook	[3]	-	26	14	37	13.5	29	16	24	15
Total enrolment	159		205		172		179		160	
The number of facets surveyed differed across the sessions	6		14		16		16		16	
Number of survey responses	63		101		63		62		21	

[1] Percentage within implementation regarding facet as moderately to extremely important to learning
Source: Student subject evaluation surveys submitted online at the end of each session.
[2] Ranking based upon proportion of students reporting aspect as moderately to extremely import to their learning. Data selected from Morris, Porter, and Griffiths, 2007, p. 161.
[3] Not surveyed
[4] Given to students as a marking guide (Source: annotated journal)
[5] Given to students as a checklist (Source: annotated journal)

Appendix 6.18

STAT131: Changing aspects of ***student evaluation surveys*** across all implementations

Aspect description	Autumn, 2003	Autumn, 2004	Spring, 2004	Autumn, 2005	Spring, 2005
Attendance and commitment patterns	No lab manual completion questions	✓	✓	✓	✓
Importance of lab tasks and solutions	×	Importance of learning strategy tasks not canvassed	✓	✓	✓
Importance of teamwork	×	Importance of teamwork not canvassed	✓	✓	✓
Importance of lab teaching	×	Importance of theory review not canvassed	✓	✓	✓
Importance of assignments	×	✓	✓	✓	✓
Exam preparedness	×	✓	✓	✓	✓
Perception of statistical learning	×	✓	✓	✓	✓
Reported assignment marks	×	✓	✓	✓	✓
Perceptions of topic learning	×	×	✓	✓	✓
Perceptions of graduate attribute achievement	×	×	×	With particular reference to the context of statistical thinking	With particular reference to the context of statistical thinking
Domestic/international student status	×	×	×	✓	✓
Gender	×	×	×	✓	✓

Source: Student survey data

Appendix 6.19

STAT131: Proportions of students responding as **moderately confident or confident in topic learning** across surveyed implementations

Topic	Spring, 2004		Autumn, 2005		Spring, 2005	
	Percentage[1]	Ranking[2]	Percentage[1]	Ranking[2]	Percentage[1]	Ranking[2]
SPSS	83	2	61	3	81	1.5
Data exploration	94	1	81	1	81	1.5
Correlation	73	3	69	2	72	3
Binomial and Poisson distributions	62	6.5	55	4	38	10
Model fit	65	4	53	5	62	4
Normal and Exponential distributions	64	5	44	7	43	8.5
Confidence intervals	57	8.5	40	9	52	5
Hypothesis tests	57	8.5	48	6	43	8.5
Markov chains	46	10	36	10	48	6
Formulae selection	62	6.5	42	8	48	7
Number of survey responses	63		62		21	
Total enrolment	172		179		160	

[1] Source: Student survey data

[2] Ranked by percentage reporting confidence

Appendix 6.20

STAT131: Percentages of students believing they **made progress toward achievement of graduate attributes**[3] across surveyed implementations

Graduate attribute	Autumn, 2005		Spring, 2005	
	Percentage[1]	Ranking[2]	Percentage[1]	Ranking[2]
Critical thinking	43	7	48	3.5
Communication of ideas and concepts	47	6	43	5
Teamwork	58	2	57	1
Evidence-based evaluation	53	4	52	2
Use of technology	63	1	48	3.5
Analysis	55	3	38	6
Personal responsibility for learning	48	5	33	7
Number of survey responses	62		21	
Total enrolment	179		160	

[1] Source: Student survey data

[2] Ranked by percentage with no significant correlation between the ranks of the two sessions

[3] Amalgamation of responses selecting ‘some progress’ and ‘a great deal of progress’

Appendix 6.21

STAT131: Descriptives[1] for all assessment across all implementations

Assessment task[5]	Autumn, 2003			Autumn, 2004			Spring, 2004			Autumn, 2005			Spring, 2005		
	N	Mean	S.D[2]	N	Mean	S.D[2]	N	Mean	S.D[2]	N	Mean	S.D[2]	N	Mean	S.D[2]
Final mark[4]	143	54.9	12.6	201	61.9	14.9	156	65.7	16.6	150	59.0	16.4	139	57.3	16.7
Final exam /100[4]	143	39.1	12.2	201	52. 9	18.9	156	60.4	19.2	150	54.4	18.4	139	51.0	19.0
Presentation /5	143	4.5	0.6	192	4.3	0.7	148	4.6	0.5	142	4.2	0.6	126	4.5	0.4
Assignment 1 /10	157	6.1	1.8	200	6.3	1.7	151	5.9	1.6	159	5.3	1.9	136	4.9	1.7
Assignment 2 /10	151	8.6	1.2	196	6.9	2.0	154	7.7	1.4	146	6.2	2.0	132	6.0	1.8
Assignment 3 /10	145	7.5	1.7	193	6.8	2.0	144	8.1	1.7	136	5.6	2.2	124	7.3	2.0
Midterm /10	152	6.1	2.2	197	5.5	1.9	154	6.1	1.8	142	5.6	2.0	141	5.2	2.3
Summary /5	128	3. 6	0.9	[3]	[3]	[3]	[3]	[3]	[3]	[3]	[3]	[3]	[3]	[3]	[3]
Laboratory mark /10	[3]	[3]	[3]	205	9.4	1.6	153	8.8	1.6	174	8.7	2.8	136	9.2	1.5

[1] Source: Student subject assessment data maintained by subject coordinator

[2] Standard deviation.

[3] Not assessed this session

[4] Zero have been removed as they represent non completion in the subject and not necessarily low achievement

[5] Zero marks have been removed from all assessment items as they represent non submission and not necessarily low achievement

[6] Median laboratory marks (reflecting completion of tasks) was 10 for all sessions except spring 2004 when the median was 9.

Appendix 6.22

Final exam marks: Significant post hoc comparisons (differences checked with ✓) of paired means for final marks (**equal variances not assumed** - Tamhane and Dunnett T3)

	Autumn, 2003	Autumn, 2004	Spring, 2004	Autumn, 2005	Spring, 2005
Autumn, 2003		✓	✓	✓	✓
Autumn, 2004	✓		✓		
Spring, 2004	✓	✓			✓
Autumn, 2005	✓				
Spring, 2005	✓		✓		

Note: The mean for spring session in 2004 was significantly different from all sessions. Assessment for the spring 2004 session closely mirrored that for autumn.

Appendix 6.23

Final exam marks: Significant post hoc comparisons (differences checked with ✓) of paired means for final marks (**equal variances assumed** - Bonferroni)

	Autumn, 2003	Autumn, 2004	Spring, 2004	Autumn, 2005	Spring, 2005
Autumn, 2003		✓	✓	✓	✓
Autumn, 2004	✓		✓		
Spring, 2004	✓	✓		✓	✓
Autumn, 2005	✓		✓		
Spring, 2005	✓		✓		

Note: The mean for the spring session in 2004 was significantly different from all sessions. Assessment for the spring 2004 session closely mirrored that for autumn.

Appendix 6.24: Percentages of students reporting progress[1] in achievement of *graduate qualities* by *thinking statistically*

	Attribute	Autumn 2005[2]	Spring 2005
Developed ability to examine problems in context and the assumptions underlying statistical analysis.	Informed problem solver	43	48
Developed ability to write meaningful paragraphs containing statistical concepts and reasoning.	Effective communicator	47	43
Increased capacity for and understanding of teamwork through working locally and/or at a distance in supporting personal learning and that of others.	Effective responsible communicator	58	57
Developed ability to make choices in the analysis of data and logically justify these choices.	Informed problem solver	53	52
Enhanced ability to use technology to analyse, organise and present data as information.	Informed independent learner	63	48
Enhanced ability to explore data in order to present improved solutions.	Informed problem solver	55	38
Recognition of responsibilities and obligations as a student and researcher.	Responsibility	48	33
N		62	21

[1] Amalgamation of responses selecting ‘some progress’ and ‘a great deal of progress’

[2] Not surveyed prior to autumn 2005

Appendix 6.25: A Sample student survey used in spring 2005

Change Evaluation STAT131 SPRING 2005

The primary purpose of this survey is to provide me with information that can assist in the development of the subject so students can complete it in different ways. Some students attend lectures others choose to use WebCT and some use both. Some resources are more useful than others. The feedback of ALL students is valuable in this process. Note: summaries of this data may be used in research publications.

Question 1
Do you consider that you have completed the lecture component by

a. Attending virtually all lectures/
b. Working through the online lectures on a regular weekly -fortnightly basis?
c. Working through the lectures as they are required for assessment?
d. Attending the lectures and working through the online lectures?
e. Other?

Question 2
How many weeks would you have attended the Monday STAT131 lectures?

a. Did not attend Monday lectures
b. 1-3 weeks
c. 4-6 weeks
d. 7-9 weeks
e. 10-13 weeks

Question 3
How many weeks did you attend the STAT131 Thursday lecture?

a. Did not attend the Thursday lecture
b. 1-3 weeks
c. 4-6 weeks
d. 7-9 weeks
e. 10-13 weeks

Question 4
How did you work through the laboratory manual?

a. Downloaded the worked solutions to most tasks

b. Started completing the tasks and downloaded to complete them

c. Essentially completed all laboratory tasks

d. Essentially completed all laboratory tasks and checked them from the worked solutions

Question 5

Including lectures and labs and work outside class, on average per week how much time did you spend for the first twelve weeks of session working on STAT131?

a. 0-5 hours per week

b. 6-8 hours per week

c. 9-11 per week

d. 12-14 hours per week

e. 15-17 hours per week

f. More than 17 hours per week

Question 6

How important were the lectures for helping you understand and learn statistics?

a. Not applicable - I rarely went to lectures

b. I learned and understood very little from the lectures I attended

c. Lectures were moderately useful for learning and understanding in this subject

d. Lectures were extremely important for me in learning and understanding statistics

Question 7

How important were the lecture notes for helping you understand and learn statistics?

a. Not applicable - I rarely referred to lecture notes

b. I learned and understood very little from the lectures notes I read

c. Lecture notes were moderately useful for learning and understanding in this subject

d. Lecture notes were extremely important for me in learning and understanding statistics

Question 8

How useful the textbook UTTS for helping you understand and learn statistics?

a. I did not really look at the textbook

b. The textbook were of little importance in helping me understand and learn the STAT131 subject

c. The textbook moderately important in helping me understand and learn the STAT131 subject

d. The textbook UTTS was extremely important in helping me understand and learn the STAT131 material

Question 9

How important were the online PowerPoint lectures for helping you understand and learn statistics?

a. Not applicable - I rarely consulted the online lectures

b. I learned very little from the online lectures I read

c. Online lectures were moderately useful for learning in this subject

d. Online lectures were extremely important for me in learning statistics

Question 10

How important were the laboratory classes for helping you understand and learn statistics?

a. I rarely attended laboratory class

b. Laboratory class was of little importance in helping me understand and learn the STAT131 subject

c. Laboratory class was moderately important in helping me understand and learn the STAT131 subject

d. Laboratory class was extremely important in helping me understand and learn the STAT131 subject

Question 11

How important was the tutor in laboratory classes for helping you understand and learn statistics?

a. I rarely attended lab classes

b. The tutor rarely helped in STAT131 lab classes

c. The tutor spoke too much and so we could not complete enough lab tasks

d. The tutor was moderately helpful in lab class

e. The tutor was extremely helpful in laboratory classes

Question 12

How important is the laboratory manual as a resource helping you understand and learn statistics?

a. I did not really use the laboratory manual

b. The laboratory manual is of little importance as a resource in helping me understand and learn the STAT131 subject

c. The laboratory manual is moderately important as a resource in helping me understand and learn the STAT131 material

d. The laboratory manual is extremely important as a resource in helping me understand and learn the STAT131 material

Question 13

How important were the laboratory tasks for helping you understand and learn statistics?

a. I rarely did laboratory tasks
b. Laboratory tasks were of little importance in helping me understand and learn the STAT131 subject
c. Laboratory tasks were moderately important in helping me understand and learn the STAT131 material
d. Laboratory tasks were extremely important in helping me understand and learn the STAT131 material

Question 14
How important were the worked solutions for lab tasks, midterms and exams for helping you understand and learn statistics?

a. I rarely looked at worked solutions
b. Worked solutions were of little importance in helping me understand and learn the STAT131 subject
c. Worked solutions were moderately important in helping me understand and learn the STAT131 material
d. Worked solutions were extremely important in helping me understand and learn the STAT131 material

Question 15
How important were the exercises using concept maps, frames and other strategies for helping you understand and learn statistics?

a. I did not really use the these techniques
b. These techniques were of little importance in helping me understand and learn the STAT131 subject
c. These techniques were moderately important in helping me understand and learn the STAT131 material
d. These techniques were extremely important in helping me understand and learn the STAT131 material

Question 16
How important were the objectives listed in the laboratory manual for helping you understand and learn statistics?

a. I did not really look at the objectives
b. The objectives were of little importance in helping me understand and learn the STAT131 subject
c. The objectives were moderately important in helping me understand and learn the STAT131 subject
d. The objectives were extremely important in helping me understand and learn the STAT131 material

Question 17
How important were the assignments for helping you understand and learn statistics?

a. I did little of the assignment work myself
b. The assignments were of little importance in helping me understand and learn the STAT131 subject
c. The assignments were moderately important in helping me understand and learn the STAT131 subject
d. The assignments were extremely important in helping me understand and learn the STAT131 material

Question 18

How important were the marking guidelines/self check listed on assignments for helping you understand and learn statistics?

a. I did not really look at the marking guidelines/self check
b. The marking guidelines/self check were of little importance in helping me understand and learn the STAT131 subject
c. The marking guidelines/self check were moderately important in helping me understand and learn the STAT131 subject
d. The marking guidelines/self check were extremely important in helping me understand and learn the STAT131 material

Question 19

How important was the midterm exam for helping you understand and learn statistics?

a. I did not really prepare for the midterm
b. The midterm was of little importance in helping me understand and learn the STAT131 subject
c. The midterm was moderately important in helping me understand and learn the STAT131 material
d. The midterm were extremely important in helping me understand and learn the STAT131 material

Question 20

How important was the forum for helping you understand and learn statistics?

a. I did not really use the forum
b. The forum was of little importance in helping me understand and learn the STAT131 subject
c. The forum was moderately important in helping me understand and learn the STAT131 material
d. The forum were extremely important in helping me understand and learn the STAT131 material

Question 21

How important was the teamwork and group work for helping you understand and learn statistics?

a. The rarely engaged in teamwork or group work

b. The teamwork/ group work was of little importance in helping me understand and learn the STAT131 subject

c. The teamwork/group work was moderately important in helping me understand and learn the STAT131 material

d. The teamwork/group work was extremely important in helping me understand and learn the STAT131 material

Question 22

How did you find the use of teams?

a. I work better on my own

b. We found it difficult to work together and basically completed the tasks individually

c. One of the team did more work than the other

d. We found it difficult to work together but continued

e. We worked well together and found it useful

Question 23

In lab class theory review was …

a. There was too little theory review to help with the tasks

b. The amount of theory covered was well balanced with time left to work on tasks

c. There was too much theory review and not enough emphasis on the tasks

Question 24

If I complete all lab classes, rather than just copy and paste solutions, I believe that I will …

a. Not have any time to study. I still have too much to do.

b. Still need to thoroughly review all lecture material and write summary notes as I don't know what to expect in the exam

c. Need only basic revision for the exam

d. Be well prepared for the exam

Question 25

In this subject on statistics, I believe that …

a. The subject is too difficult

b. I have tried to make sense of difficult material and have limited success

c. I have tried to make sense of difficult material but I feel that I have been unsuccessful

d. I have tried to make sense of difficult material, and have a moderate amount of success

e. I have learned a great deal

Question 26

Compared to when I started STAT131, I now view …

a. Statistics as less relevant to my life and/or anticipated profession

b. About the same relevance as when I commenced STAT131

c. Statistics as more relevant to my life and/or anticipated profession

Question 27

What mark did you get your first assignment out of 50?

Question 28

What mark did you get your second assignment out of 65?

Question 29

What mark did you get for the midterm out of 25?

Question 30

How comfortable are you now that you could set up a data set in SPSS and produce output?

a. Not at all

b. Might have a little difficulty

c. Moderately confident

d. Could do this

Question 31

How comfortable are you now that you could do basic exploration identifying and describing the centre, shape, spread and outliers of data?

a. Not at all

b. Might have a little difficulty

c. Moderately confident

d. Could do this

Question 32

How comfortable are you now that you could do basic plotting, identifying relationships and conducting a regression analysis on two variables?

a. Not at all

b. Might have a little difficulty

c. Moderately confident

d. Could do this

Question 33

How comfortable are you solving problems using the Binomial and Poisson distributions?

a. Not at all

b. Might have a little difficulty

c. Moderately confident

d. Could do this

Question 34

How comfortable are you now that you could determine whether a set of data fit a particular model?

a. Not atl all

b. Might have a little difficulty

c. Moderately confident

d. Could do this

Question 35

How comfortable are you solving problems using the Exponential and Normal distributions?

a. Not at all

b. Might have a little difficulty

c. Moderately confident

d. Could do this

Question 36

How comfortable are you now that you could set up and interpret confidence intervals?

a. Not at all

b. Might have a little difficulty

c. Moderately confident

d. Could do this

Question 37

How comfortable are you now that you could set up, test and interpret hypotheses tests?

a. Not at all

b. Might have a little difficulty

c. Moderately confident

d. Could do this

Question 38

How comfortable are you now that you could use a Markov Chain to calculate probabilities?

a. Not at all

b. Might have a little difficulty

c. Moderately confident

d. Could do this

Question 39

How comfortable are you now that you identify appropriate formulae for different situations?

a. Not at all

b. Might have a little difficulty

c. Moderately confident

d. Could do this

Question 40

I believe that my efforts in this subject have enabled me to examine problems in context and the assumptions underlying statistical analysis.

a. Not at all

b. I have tried but feel that I have made no progress

c. I have tried but feel I have made limited progress

d. I have tried but feel I have made some progress

e. I have tried but feel I have made a great deal of progress

Question 41

I believe that my efforts in this subject have increased my confidence in writing meaningful paragraphs containing statistical concepts and reasoning.

a. Not at all

b. I have tried but feel that I have made no progress

c. I have tried but feel I have made limited progress

d. I have tried but feel I have made some progress

e. I have tried but feel I have made a great deal of progress

Question 42

I believe that my efforts in this subject have helped me to increase my capacity for, and understanding of, teamwork by working locally and/or at a distance to support my own and the learning of others.

a. Not at all
b. I have tried but feel that I have made no progress
c. I have tried but feel I have made limited progress
d. I have tried but feel I have made some progress
e. I have tried but feel I have made a great deal of progress

Question 43

I believe that my efforts in this subject have helped me to develop my ability to make choices in the analysis of data and logically justify these choices.

a. Not at all
b. I have tried but feel that I have made no progress
c. I have tried but feel I have made limited progress
d. I have tried but feel I have made some progress
e. I have tried but feel I have made a great deal of progress

Question 44

I believe that my efforts in this subject have helped me to develop my ability to use technology to analyse, organise and present data as information.

a. Not at all
b. I have tried but feel that I have made no progress
c. I have tried but feel I have made limited progress
d. I have tried but feel I have made some progress
e. I have tried but feel I have made a great deal of progress

Question 45

I believe that my efforts in this subject have helped me to develop my ability to explore data in order to present improved solutions.

a. Not at all
b. I have tried but feel that I have made no progress
c. I have tried but feel I have made limited progress
d. I have tried but feel I have made some progress
e. I have tried but feel I have made a great deal of progress

Question 46

I believe that my efforts in this subject have helped me to recognise my responsibilities and obligations as a student and researcher.

a. Not at all
b. I have tried but feel that I have made no progress
c. I have tried but feel I have made limited progress
d. I have tried but feel I have made some progress
e. I have tried but feel I have made a great deal of progress

Question 47
Select your origin

a. I am an International student
b. I am a domestic / Australian student

Question 48
Select your gender

a. Female
b. Male

Question 49
How best can the STAT131 subject be improved?

Question 50
If you would be willing to be personally interviewed about these responses, please include your name and student number here and (subject designer) or (researcher/teacher) will be in touch. It is important that we interview students from all levels of achievement to get an unbiased perspective on how to improve the subject.

Question 51
If you are happy to have these responses linked to your final exam marks so that we can analyse student perceptions of learning with final outcomes could you please supply your student number.

Appendix 6.26: Percentage of marks allocated to questions on bivariate relationships (regression/correlation) in the autumn 2003 final exam (classified using the revised taxonomy of Bloom[1])

Please see print copy for Appendix 6.26

[1] Anderson and Krathwohl, 2001.

Notes:

1. Shaded areas indicate higher order demand (see Appendix 6.29)
2. Details of coding have been given in Appendix 6.31a

Appendix 6.27: Percentage of marks allocated to questions on bivariate relationships (regression/correlation) in the autumn 2004 final exam as classified using the revised taxonomy of Bloom[1]

Please see print copy for Appendix 6.27

[1] Anderson and Krathwohl, 2001.

Notes:

1. Shaded areas indicate higher order demand (see Appendix 6.29)

2. Details of coding have been given in Appendix 6.31b

Appendix 6.28: Percentage of marks allocated to questions on bivariate relationships (regression/correlation) in the autumn 2005 final exam as classified using the revised taxonomy of Bloom[1]

Please see print copy for Appendix 6.28

[1] Anderson and Krathwohl, 2001.

Notes:

1. Shaded areas indicate higher order demand (see Appendix 6.29)
2. Details of coding have been given in Appendix 6.31c

Appendix 6.29: Classification of knowledge types and cognitive processing skills[1]

Please see print copy for Appendix 6.29

[1] Anderson and Krathwohl, 2001.

Note: italicised elements regarded as 'higher order'

Legend for all slides on Knowledge and Skills:

Notes: increase in classification code indicates an increase in complexity/order of performance demand;

Italics indicate knowledge and skills identified as higher order by the researcher.

Appendix 6.30: Example of STAT131 Assignment 1.

Please see print copy for Appendix 6.30

Appendix 6.31a: Coding of examination questions (correlation of quantitative variables) using the revised taxonomy of Bloom* (see Appendix 6.29) - Autumn Session 2003

Question	Coding
Question 1 Note: SPSS output given (scatterplot, Correlations, Model summary (r, r^2), Coefficients (t tests)) a) What four attributes of a scatterplot are of interest?	4 marks Knowledge: Fact – detail (1.2) Cognition: Recall (1.2) Coding: 4(1.2 × 1.2)*
b) What does this scatterplot indicate about the nature of the relationship between a child's height and parent's height?	2 marks Knowledge: Fact – detail (1.2) Cognition: Understand – interpret (2.1) Coding: 2(1.2 × 2.1)*
c) What is the value of Pearson's r for these data?	1 mark Knowledge: Fact – detail (1.2) Cognition: Analyse – differentiate (4.1) Coding: (1.2 × 4.1)*
d) Explain the principle behind choosing the line of best fit in order to predict one variable from another. Use a diagram to help explain.	2 marks Diagram: 1 mark Knowledge: Procedural – Subject specific skills & algorithms (3.1) Cognition: Understand – explain (2.7) Explanation: 1 mark Knowledge: Procedural – Subject specific skills & algorithms (3.1) Cognition: Understand – interpret (2.1) Coding: (3.1 × 2.7) and ((3.1 × 2.1)*

Appendix 6.31a (continued): Coding of examination questions (correlation of quantitative variables) using the revised taxonomy of Bloom* (see Appendix 6.29) – Autumn Session 2003

Question	Question
e) What is the equation of the line of best fit through these data?	2 marks Identification of independent variable: 1 mark Knowledge: Procedural – Subject specific skills & algorithms (3.1) Cognition: Apply – implement (3.2) Constructing equation: 1 mark Knowledge: Conceptual – generalisation (2.2) Cognition: Analyse – differentiate (4.1) Coding: (3.1×3.2) and $((2.2 \times 4.1)$*
f) If a parent is 170 cm in height what height would you predict for the same sex child?	1 mark Knowledge: Procedural – skills (3.1) Cognition: Apply – execute (3.1) Coding: (3.1×3.1)*
g) Plot the line of best fit (least squares regression line) on the graph provided. (Scatterplot provided)	1 mark Knowledge: Procedural – techniques (3.2) Cognition: Understand – interpret (3.2) Coding: (3.2×3.2)*

Appendix 6.31a: Coding of examination questions (correlation of quantitative variables) using the revised taxonomy of Bloom* (see Appendix 6.29) - Autumn Session 2003

Question	Question
h) How appropriate is the line of best fit for predicting the height of the same sex child for someone drawn from this population? Explain using reference to the coefficient of determination r^2.	3 marks ‘What is the value of r^2?’: 1 mark Knowledge: Fact – detail (1.2) Cognition: Analyse - differentiate (4.1) ‘What does it measure?’: 1 mark Knowledge: Procedural – technique (3.2) Cognition: Recall (1.2) ‘Explain.’ (justify): 1 mark Knowledge: Procedural – criteria (3.3) Cognition: Evaluate – critique (5.2) Coding: (1.2×4.1) and (3.2×1.2) and (3.3×5.2)*
i) Sometimes an r^2 value in one discipline is viewed with excitement but in another discipline the same r^2 value is considered too low to be worthwhile. Why does this differ from discipline to discipline or one research project to another?	2 marks Necessity for accuracy/control: 1 mark Knowledge: Meta – strategic (4.1) Cognition: Analyse – organise (4.2) Determining what accuracy etc is required: 1 mark Knowledge: Procedural – criteria (3.3) Cognition: Understand –exemplify (2.2) Coding: (4.2×4.1) and (3.3×2.2)*

Appendix 6.31a (continued): Coding of examination questions (correlation of quantitative variables) using the revised taxonomy of Bloom* (see Appendix 6.29) - Autumn Session 2003

Question	Question
j) If the height of a parent is 120 cm, is it appropriate to predict the height of the child using the regression equation obtained from these data? Explain.	2 marks Range of values: 1 mark Knowledge: Procedural – criteria (3.3) Cognition: Analyse – differentiate (4.1) Explanation: 1 mark Knowledge: Procedural – criteria (3.3) Cognition: Analyse – organise (4.2) Coding: (3.3×4.1) and (3.3×4.2)*

Appendix 6.31b: Coding of examination questions (correlation of quantitative variables) using the revised taxonomy of Bloom (see Appendix 6.29) - Autumn Session 2004

Question	Coding
Question 2 Note: SPSS output given (scatterplot, Correlations, Model summary (r, r^2), ANOVA and Coefficients (t tests)) a) What are the two main reasons for producing a scatterplot before using Pearson's r to calculate the correlation coefficient? Explain why it is necessary to examine these features.	1 mark Knowledge: Procedural – criteria (3.3) Cognition: Understand - explain (2.7) Coding: (3.3×2.7)*
b) Write a short report on what the scatterplot for these data reveal.	3 marks Discriminating relevant detail – 2 marks Knowledge: Procedural – techniques (3.2) Cognition: Analyse - differentiate (4.1) Inferring from detail – 1mark Knowledge: Procedural – techniques (3.2) Cognition: Understand - infer (2.5) Coding: $2(3.2 \times 4.1)$ and (3.2×2.5)*
c) Under what circumstances would you leave out a point from the analysis?	2 marks Knowledge: Procedural – criteria (3.3) Cognition: Recall (1.2) Coding: $2(3.3 \times 1.2)$*

Appendix 6.31b (continued): Coding of examination questions (correlation of quantitative variables) using the revised taxonomy of Bloom (see Appendix 6.29) - Autumn Session 2004

Question	Coding
d) How would you determine the impact of a potentially influential point on Pearson's r?	1 mark Knowledge: Procedural – techniques (3.2) Cognition: Understand - explain (2.7) Coding: (3.2×2.7)*
e) What is the value of the correlation between exam marks and continuous assessment?	1 mark Knowledge: Procedural – skills (3.1) Cognition: Analyse - differentiate (4.1) Coding: (3.1×4.1)*
f) Someone described the relationship as a positive direct one but this conveys little meaning to the lay person. How else can you describe the nature of the relationship between examination marks and continuous assessment so that the description conveys meaning?	1 mark Knowledge: Procedural – techniques (3.2) Cognition: Understand - interpret (2.1) Coding: (3.2×2.1)*
g) What proportion of the variation in examination marks can be explained in terms of the assessment marks?	1 mark Knowledge: Procedural – techniques (3.2) Cognition: Analyse - differentiate (4.1) Coding: (3.2×4.1)*

Appendix 6.31b (continued): Coding of examination questions (correlation of quantitative variables) using the revised taxonomy of Bloom (see Appendix 6.29) - Autumn Session 2004

Question	Coding
h) Anne wants to use the continuous assessment marks to estimate examination marks. What is the principle behind putting a regression line through data? Illustrate your answer.	2 marks Explanation: 1 mark Knowledge: Procedural – techniques (3.2) Cognition: Understand - explain (2.7) Diagram: 1 mark Knowledge: Procedural – techniques (3.2) Cognition: Understand - interpret (2.1) Coding: (3.2×2.7) and (3.2×2.1)*
i) Using the output provided specify the equation of least squares regression line.	1 mark Knowledge: Procedural – techniques (3.2) Cognition: Analyse - differentiate (4.1) Coding: (3.2×4.1)*
j) A student missed the final examination and is unable to sit another. Use the least squares regression equation to predict his examination mark given that he got 30 in the continuous assessment.	1 mark Knowledge: Procedural – skills (3.1) Cognition: Apply - execute (3.1) Coding: (3.1×3.1)*

Appendix 6.31b (continued): Coding of examination questions (correlation of quantitative variables) using the revised taxonomy of Bloom (see Appendix 6.29) - Autumn Session 2004

Question	Coding
k) Another student needs an assessment for a missed assignment. This student has a mark from the final examination. What do you suggest that Anne do to estimate this mark?	1 mark Knowledge: Procedural – criteria (3.3) Cognition: Evaluate - checking (5.1) Coding: (3.3 × 5.1)*
l) Should she use an independent t-test or a related t-test (otherwise known as a paired t-test) to test for differences in means of continuous assessment and the examination? Explain why she would do this.	1 mark Knowledge: Procedural – criteria (3.3) Cognition: Evaluate - checking (5.1) Coding: (3.3 × 5.1)*
m) What two plots would you recommend she use to highlight the differences in marks? Sketch the plots, without reference to the data.	2 marks Specification of plots: 1 mark Knowledge: Procedural – techniques (3.2) Cognition: Recall (1.2) Sketching the plots: 1 mark Knowledge: Procedural – techniques (3.2) Cognition: Analyse - differentiate (4.1) Coding: (3.2 × 1.2) and (3.2 × 4.1)*
n) What statistical assumptions must she consider when conducting a t-test?	2 marks Knowledge: Procedural – criteria (3.3) Cognition: Recall (1.2) Coding: 2(3.3 × 1.2)*

Appendix 6.31c: Coding of examination questions (correlation of quantitative variables) using the revised taxonomy of Bloom (see Appendix 6.29) - Autumn Session 2005

Notes:

1. This question was marked out 30 marks and hence was scaled for comparison with 2003 and 2004 exam questions
2. These questions were more textually framed than in 2003 and 2004.

Question	Coding
Question 1 Note: SPSS output given (scatterplot, correlations, Model summary (r, r^2), ANOVA and Coefficients (t tests)) Part A	
a) If we were doing a study of glucose levels comparing males and females and we were given a data set, what information would we seek from the literature to provide context for our data?	1 mark Knowledge: Conceptual - theories (2.3) Cognition: Create - generate (6.1) Coding: (2.3×6.1)*
b) If you were at the phase of checking your data for accuracy, how might you use this scatterplot?	1 mark Knowledge: Procedural - techniques (3.2) Cognition: Recall (1.2) Coding: (3.2×1.2)*

Appendix 6.31c (continued): Coding of examination questions (correlation of quantitative variables) using the revised taxonomy of Bloom (see Appendix 6.29) - Autumn Session 2005

Question	Coding
c) When we ask a question about relationships, we produce a scatterplot and use the scatterplot to examine five major features of the data. What are these five features and what does the scatterplot reveal?	5 marks Distinguishing the features: 3 marks Knowledge: Procedural - skills (3.1) Cognition: Analyse - distinguish (4.1) Inferring from the plot: 2 marks Knowledge: Procedural - skills (3.1) Cognition: Understand - infer (2.5) Coding: $3(3.1 \times 4.1)$ and $2(3.1 \times 2.5)$*
d) Using the scatterplot provided, describe in plain English the nature of the relationship between glucose0 and glucose3.	1 mark Knowledge: Procedural - skills (3.1) Cognition: Understand – interpret (2.1) Coding: (3.1×2.1)*
e) Based on the data in the scatterplot, is it appropriate to use Pearson's r as a measure of correlation for this data? Explain why or why not.	2 marks Evidence of non-linearity: 1 mark Knowledge: Procedural - skills (3.1) Cognition: Understand – interpret (2.1) Scatter/extreme values: 1 mark Knowledge: Procedural - skills (3.1) Cognition: Understand – explain (2.7) Coding: (3.1×2.1) and (3.1×2.7)*

Appendix 6.31c (continued): Coding of examination questions (correlation of quantitative variables) using the revised taxonomy of Bloom (see Appendix 6.29) - Autumn Session 2005

Question	Coding
Part B a) Of what does Pearson's r provide a measure?	1 mark Knowledge: Procedural - techniques (3.2) Cognition: Understand – explain (2.7) Coding: (3.2×2.7)*
b) What is the value of Pearson's r for these data?	1 mark Knowledge: Procedural - skills (3.1) Cognition: Analyse – differentiate (4.1) Coding: (3.1×4.1)*
c) What does Pearson's r suggest about the relationship between glucose0 and glucose3?	2 marks Retrieving value: 1 mark Knowledge: Procedural - skills (3.1) Cognition: Analyse – differentiate (4.1) Inferring from value: 1 mark Knowledge: Procedural - techniques (3.2) Cognition: Understand – infer (2.5) Coding: (3.1×4.1) and (3.2×2.5)*
Part C a) Reading from the output, what is the equation of the line of best fit through these data?	1 mark Knowledge: Procedural - techniques (3.2) Cognition: Analyse - differentiate (4.1) Coding: (3.2×4.1)*

Appendix 6.31c (continued): Coding of examination questions (correlation of quantitative variables) using the revised taxonomy of Bloom (see Appendix 6.29) - Autumn Session 2005

Question	Coding
b) Given an initial glucose reading of 10mmol/L, what is the predicted value of glucose3?	1 mark Knowledge: Procedural - skills (3.1) Cognition: Apply - execute (3.1) Coding: (3.1×3.1)*
c) Is it appropriate to make the prediction in b)? Explain using reference to 1. the proportion of variation in glucose3 explained by glucose0 2. and, from the Analysis of variance table, Sig.	2 marks Reference to r^2: 1 mark Knowledge: Procedural - techniques (3.2) Cognition: Understand - explain (2.7) Reference to p value: 1 mark Knowledge: Procedural - techniques (3.2) Cognition: Understand - infer (2.5) Coding: (3.2×2.7) and (3.2×2.5)*
d) Given an initial glucose value of 50 mmol/L what is the predicted value of glucose3?	1 mark Knowledge: Procedural - skills (3.1) Cognition: Apply - execute (3.1) Coding: (3.1×3.1)*
e) You advise the researcher that the prediction made in d) is not appropriate. Explain why it is not appropriate.	1 mark Knowledge: Procedural - criteria (3.3) Cognition: Understand - infer (2.5) Coding: (3.3×2.5)*

Appendix 7.1a: Subject presentation (***lectures)*** before and after implemented changes with associated student survey responses (n=150)			
Subject Presentation	Pre- implementation	Post implementation	Associated student survey responses
Lectures	Expository Large group (223) across multiple campuses Little or no interaction Non compulsory but attendance strongly recommended	No change	Only 40% reported regular (11-13 weeks) Only 68% perceived lectures as moderately to extremely important to their learning Only two students commented specifically on lectures: • link between lectures and tutorials to be made more transparent • more practical problems - less theory (*return to surface learning*)

Appendix 7.1b: Subject presentation (***tutorials***) before and after implemented changes with associated student survey responses (n=150)			
Subject Presentation	Pre- implementation	Post-implementation	Associated student survey responses
Tutorials	25-30 students; Compulsory attendance; Group discussions; Collaborative learning encouraged; Submission of tasks completed before tutorial class; Tutor's role not defined: supervisory/participative; Participation marks based on attendance and submission of tasks but not defined prescriptively.	25-30 students; Compulsory attendance; Group discussions; Collaborative learning promoted: • Group presentation of tutorial tasks; • Peer evaluations of group presentations using the marking criteria for the presentation assessment task. Compilation of *portfolio* from: • tutorial tasks and solutions; • group evaluations; • copies of exemplars from the web; • weekly reflection on their learning in the topic additional relevant support material. Students used their portfolios as exam resources; Tutor's role more defined: • facilitated group work and presentations; • marked presentations using the students' criteria; and	91% reported regular (11-13 weeks) attendance Percentage perceiving aspect as moderately to extremely important to their learning: • Tutorials: 85% • Tutor: 84% • Tutorial exercises: 81% • Presentation exercises: 67% • Group work: 71% • Portfolio: 82% 74% of students report a positive experience of group work Student comments Two negative comments about the tutor (*no teaching*); Many students failed to perceive the relevance of the presentation exercises; Requests for *directed questions* on the articles to facilitate discussion; Many requests for practical text book type questions/teaching approach to tutorial (*return to surface learning which is more comfortable but less beneficial to deeper*

Appendix 7.1b: Subject presentation (***tutorials***) before and after implemented changes with associated student survey responses (n=150)			
Subject Presentation	Pre- implementation	Post-implementation	Associated student survey responses
		• checked attendance; • marked portfolios for completion. Participation marks based on attendance and portfolio completion but still not defined prescriptively	*learning*); Some comments highlighted inconsistencies in tutor approaches (e.g. presentations *before* group discussions) (*Commitment of tutors essential to successful implementation*); Many comments indicated students' failure to appreciate the reasons for the inclusion of analyses of journal articles in the exercises. *This indicates that there needs to be more explicit statement of the learning objectives for this subject for the students.*

Appendix 7.1c: Subject presentation (***web resources***) before and after implemented changes with associated student survey responses (n=150)			
Subject Presentation	Pre- implementation	Post-implementation	Associated student survey responses
Lecture notes	PowerPoint and PDF files available Shared authorship **Usually** available before lectures	No change	70% of students perceived lecture notes as moderately to extremely important to their learning Many students expressed concern that notes were not always available before lectures. (*Further supports teaching fear that students use notes in lieu of attendance*)
Online Resources	Notes and support material Available when required	No change	69% of students perceived online resources as moderately to extremely important to their learning
Assessment	Available ahead of due dates	No change	Percentage perceiving aspect of assessment regimen as moderately to extremely important to their learning: • Essay: 83% • Essay outline: 75% • Presentation: 75% • Marking criteria: 67% • Peer evaluation: 53% • Portfolios: 82% *Need to survey perception of reflection exercises and fairness of assessment*
Supplementary	Some sample solutions to practical problems	Some exemplars available for complex	Several students requested more worked

Appendix 7.1c: Subject presentation (***web resources***) before and after implemented changes with associated student survey responses (n=150)			
Subject Presentation	Pre- implementation	Post-implementation	Associated student survey responses
Resources	available for reference only	problems	examples (*return to surface learning*) Student comments Two negative comments about the tutor (*no teaching*) Many students failed to perceive the relevance of the presentation exercises Requests for *directed questions* on the articles to facilitate discussion Many requests for practical text book type questions/teaching approach to tutorial (*return to surface learning which is more comfortable but less beneficial to deeper learning*)

Appendix 7.2: Subject assessment before and after implemented changes with associated student mark summaries (N=223)

Task	Pre-Implementation	Post-Implementation	Student Performance
Essay	Individual submission; Based on group research for group presentation (allowing collaboration on content) Submitted near end of session; No specification of learning outcomes; High standard of academic writing expected.	Individual submission; Based on group research for group presentation (allowing collaboration on content); Submitted near end of session; Specification of learning outcomes aligned with the of marking criteria; High standard of academic writing expected supported by detail in the marking criteria; Familiar criteria: the same as for peer evaluations in tutorials and parallel to essay criteria.	Maximum possible: 20 marks Descriptives: • Mean= 12.3 • Median= 12.5 • SD= 2.9 • IQR= 3.8
Presentation	Group presentation; Last week of session; **No** detailed marking criteria specified (either for staff or students); Potential for subjectivity; Marking included presentation skills, content and logical structure; Team skills not evaluated.	Group presentation; Last week of session; Detailed marking criteria specified (for both staff and students); Familiar criteria: the same as for peer evaluations in tutorials and parallel to essay criteria; Less subjectivity in marking; Marking included presentation skills, content and logical structure; Team skills still not evaluated.	Maximum possible: 15 marks Descriptives: • Mean= 10.9 • Median= 11.0 • SD= 1.4 • IQR= 2.0

Appendix 7.2 (continued): Subject assessment before and after implemented changes with associated student mark summaries (N=223)			
Task	Pre-Implementation	Post-Implementation	Student Performance
Participation	Marks allocated equally for attendance and completion of the set tasks but guidelines for marking were verbal only	Marks allocated equally for contribution to class discussion and completion of the portfolio but again guidelines for marking were verbal only	Maximum possible: 10 marks Descriptives: ♦ Mean= 6.3 ♦ Median= 6.5 ♦ SD= 1.2 ♦ IQR= 1.0
Exam	Selection of three out of five essay questions related to the weekly topics; One major accounting problem with an accompanying discussion/argument.	Selection of three out of five essay questions related to the weekly topics; One major accounting problem with an accompanying discussion/argument; Use of portfolio as an exam resource.	Descriptives (%): • Mean= 48.3% • Median= 48.5% • SD= 9.4% • IQR= 12.5% Essay questions (select 3 from 5) (Maximum each question: 20)* 1: Mean= 9.4 SD= 3.7 2: Mean= 9.7 SD= 2.9 3: Mean= 8.5 SD= 3.1 4: Mean= 8.7 SD= 2.7 5: Mean= 10.2 SD= 2.5 Part B (maximum possible: 40 marks) Mean= 20.7 SD= 7.0

Appendix 7.3: Student subject evaluation survey

Subject Evaluation
Accounting Theories and Philosophies

The primary purpose of this survey is to provide information which might be used to assist in the development of this subject so that students can achieve maximum benefit from its presentation. It surveys your perception of the course components in terms of their relevance to your learning and your confidence in your learning in the specified topic areas.

Please select the most appropriate responses from the given alternatives.

Question 1
Your attendance at lectures is best represented by which of the following statements:

A. Attending lectures for between 11-13 weeks
B. Attending lectures for between 6-10 weeks
C. Attending lectures for less than 6 weeks
D. I did not attend lectures, but believe that I secured the relevant information by other means.

Question 2
Your attendance at tutorials is best represented by which of the following statements:

A. Attending tutorials for between 11-13 weeks
B. Attending tutorials for between 6-10 weeks
C. Attending tutorials for less than 6 weeks
D. I did not tutorials, but believe that I secured the relevant information by other means.

Question 3
Including lectures, tutorials and preparation/study/assignments outside of class, how much time do you believe that you spent working on this subject on average per week?

A. 15 hours or more
B. 10-15 hours
C. 5-10 hours
D. Less than 5 hours.

Question 4
How important were the lectures for helping your learning in Accounting Theories and Philosophies?

A. Lectures were extremely important to my learning
B. Lectures were moderately important to my learning
C. Lectures were not very helpful for my learning
D. Not applicable as I rarely attended lectures.

Question 5

How important were the lecture notes for helping your learning in Accounting Theories and Philosophies?

A. Lecture notes were extremely important to my learning
B. Lecture notes were moderately important to my learning
C. Lecture notes were not very helpful for my learning
D. Not applicable as I rarely/never referred to lecture notes.

Question 6
How important was the text book for helping your learning in Accounting Theories and Philosophies?

A. The text book was extremely important to my learning
B. The text book was moderately important to my learning
C. The text book was not very helpful for my learning
D. Not applicable as I rarely/never referred to the text book.

Question 7
How important were the online resources for helping your learning in Accounting Theories and Philosophies?

A. The online resources were extremely important to my learning
B. The online resources were moderately important to my learning
C. The online resources were not very helpful for my learning
D. Not applicable as I rarely/never referred to the online resources.

Question 8
How important were the tutorials for helping your learning in Accounting Theories and Philosophies?

A. The tutorials were extremely important to my learning
B. The tutorials were moderately important to my learning
C. The tutorials were not very helpful for my learning
D. Not applicable as I rarely/never attended the tutorials.

Question 9
How important were the tutorials exercises for helping your learning in Accounting Theories and Philosophies?

A. The tutorial exercises were extremely important to my learning
B. The tutorial exercises were moderately important to my learning
C. The tutorial exercises were not very helpful for my learning
D. Not applicable as I rarely/never participated in the tutorials.

Question 10
How important was your ‘portfolio’ for helping your learning in Accounting Theories and Philosophies?

A. The portfolio was extremely important to my learning
B. The portfolio was moderately important to my learning

C. The portfolio was not very helpful for my learning
D. Not applicable as I rarely/never worked on the portfolio.

Question 11
How important was the role of the tutor for helping your learning in Accounting Theories and Philosophies?

A. The tutor's role was extremely important to my learning
B. The tutor's role was moderately important to my learning
C. The tutor's role was not very helpful for my learning
D. Not applicable as the tutor did not participate in 'teaching'.

Question 12
How important were the 'Essay outline' exercises for helping your learning in Accounting Theories and Philosophies?

A. The 'Essay outline' exercises were extremely important to my learning
B. The 'Essay outline' exercises were moderately important to my learning
C. The 'Essay outline' exercises were not very helpful for my learning
D. Not applicable as I rarely/never completed the 'Essay outline' exercises.

Question 13
How important were the 'Presentation' exercises for helping your learning in Accounting Theories and Philosophies?

A. The 'Presentation' exercises were extremely important to my learning
B. The 'Presentation' exercises were moderately important to my learning
C. The 'Presentation' exercises were not very helpful for my learning
D. Not applicable as I rarely/never completed the 'Presentation' exercises.

Question 14
How important was the essay for helping your learning in Accounting Theories and Philosophies?

A. The essay was extremely important to my learning
B. The essay was moderately important to my learning
C. The essay was not very helpful for my learning
D. Not applicable as I never submitted the essay.

Question 15
How important was the presentation for helping your learning in Accounting Theories and Philosophies?

A. The presentation was extremely important to my learning
B. The presentation was moderately important to my learning
C. The presentation was not very helpful for my learning
D. Not applicable as I did little to participate in the presentation.

Question 16

How important was the group work for helping your learning in Accounting Theories and Philosophies?

A. The group work was extremely important to my learning
B. The group work was moderately important to my learning
C. The group work was not very helpful for my learning
D. Not applicable as I rarely/never participated in group work.

Question 17
How important was the group evaluation for helping your learning in Accounting Theories and Philosophies?

A. The group evaluation was extremely important to my learning
B. The group evaluation was moderately important to my learning
C. The group evaluation was not very helpful for my learning
D. Not applicable as I rarely/never participated in group evaluation.

Question 18
How important was the marking criteria for helping your learning in Accounting Theories and Philosophies?

A. The marking criteria was extremely important to my learning
B. The marking criteria was moderately important to my learning
C. The marking criteria was not very helpful for my learning
D. Not applicable as I rarely/never learned anything from the marking criteria.

Question 19
Choose the statement which best reflects your overall level of learning in Accounting Theories and Philosophies.

A. I have achieved most of the objectives of this subject
B. I found this subject challenging but feel that I have achieved most of the objectives of this subject
C. I found this subject difficult and feel that I have not achieved many of the objectives of this subject
D. I found this subject difficult and feel that I have not achieved the objectives of this subject.

For questions 20-30 please choose the statement which best reflects your level of learning in the given areas

Question 20
Management accounting theories

A. I understood this topic and feel very confident about answering questions in this area
B. I found this topic challenging but feel moderately confident about answering questions in this area
C. I found this topic difficult and do not feel confident about answering questions in this area

D. I found this topic too difficult and could not answer questions in this area.

Question 21
Regulatory theories

A. I understood this topic and feel very confident about answering questions in this area
B. I found this topic challenging but feel moderately confident about answering questions in this area
C. I found this topic difficult and do not feel confident about answering questions in this area
D. I found this topic too difficult and could not answer questions in this area.

Question 22
Accounting models HC and CPP

A. I understood this topic and feel very confident about answering questions in this area
B. I found this topic challenging but feel moderately confident about answering questions in this area
C. I found this topic difficult and do not feel confident about answering questions in this area
D. I found this topic too difficult and could not answer questions in this area.

Question 23
Accounting models CCA and CoCoA

A. I understood this topic and feel very confident about answering questions in this area
B. I found this topic challenging but feel moderately confident about answering questions in this area
C. I found this topic difficult and do not feel confident about answering questions in this area
D. I found this topic too difficult and could not answer questions in this area.

Question 24
Critical analysis of the conceptual framework & harmonisation

A. I understood this topic and feel very confident about answering questions in this area
B. I found this topic challenging but feel moderately confident about answering questions in this area
C. I found this topic difficult and do not feel confident about answering questions in this area
D. I found this topic too difficult and could not answer questions in this area.

Question 25
Cash flow accounting

A. I understood this topic and feel very confident about answering questions in this area
B. I found this topic challenging but feel moderately confident about answering questions in this area
C. I found this topic difficult and do not feel confident about answering questions in this area
D. I found this topic too difficult and could not answer questions in this area.

Question 26
Agency and Positive Accounting Theory

A. I understood this topic and feel very confident about answering questions in this area
B. I found this topic challenging but feel moderately confident about answering questions in this area
C. I found this topic difficult and do not feel confident about answering questions in this area
D. I found this topic too difficult and could not answer questions in this area.

Question 27
System Oriented Theories

A. I understood this topic and feel very confident about answering questions in this area
B. I found this topic challenging but feel moderately confident about answering questions in this area
C. I found this topic difficult and do not feel confident about answering questions in this area
D. I found this topic too difficult and could not answer questions in this area.

Question 28
Capital Market Theories

A. I understood this topic and feel very confident about answering questions in this area
B. I found this topic challenging but feel moderately confident about answering questions in this area
C. I found this topic difficult and do not feel confident about answering questions in this area
D. I found this topic too difficult and could not answer questions in this area.

Question 29
Behavioural Theory

A. I understood this topic and feel very confident about answering questions in this area
B. I found this topic challenging but feel moderately confident about answering questions in this area
C. I found this topic difficult and do not feel confident about answering questions in this area

D. I found this topic too difficult and could not answer questions in this area.

Question 30
Critical Theory

A. I understood this topic and feel very confident about answering questions in this area
B. I found this topic challenging but feel moderately confident about answering questions in this area
C. I found this topic difficult and do not feel confident about answering questions in this area
D. I found this topic too difficult and could not answer questions in this area.

For questions 31-38, choose the statement which best reflects your level of achievement of the given subject objectives.

Question 31
"… to have developed an appreciation of accounting theory including the types of theories and different research methods, evaluating the underlying assumptions, objectives, paradigms, and logic and knowledge claims of each"

A. I believe that I am extremely competent in this area
B. I believe that I am moderately competent in this area
C. I find this skill difficult and have not achieved as much as I would like in this area
D. I have not achieved this skill.

Question 32
"… to be able to identify the essential elements in the various accounting models available"

A. believe that I am extremely competent in this area
B. I believe that I am moderately competent in this area
C. I find this skill difficult and have not achieved as much as I would like in this area
D. I have not achieved this skill.

Question 33
"… to be able to demonstrate an understanding of current value income measurement models of accounting developed to date including the impact of inflation on income and value determination"

A. I believe that I am extremely competent in this area
B. I believe that I am moderately competent in this area
C. I find this skill difficult and have not achieved as much as I would like in this area
D. I have not achieved this skill.

Question 34

"… to be able to discuss and evaluate the strengths and weaknesses of traditional accounting systems"

A. I believe that I am extremely competent in this area
B. I believe that I am moderately competent in this area
C. I find this skill difficult and have not achieved as much as I would like in this area
D. I have not achieved this skill.

Question 35

"… to develop an improved ability to critically appraise or evaluate ideas"

A. I believe that I am extremely competent in this area
B. I believe that I am moderately competent in this area
C. I find this skill difficult and have not achieved as much as I would like in this area
D. I have not achieved this skill.

Question 36

"… to be able to appreciate the anthropological view of accounting"

A. I believe that I am extremely competent in this area
B. I believe that I am moderately competent in this area
C. I find this skill difficult and have not achieved as much as I would like in this area
D. I have not achieved this skill.

Question 37

"… to be able to make an effective presentation"

A. I believe that I am extremely competent in this area
B. I believe that I am moderately competent in this area
C. I find this skill difficult and have not achieved as much as I would like in this area
D. I have not achieved this skill.

Question 38

"… to be able to work more effectively in small groups"

A. I believe that I am extremely competent in this area
B. I believe that I am moderately competent in this area
C. I find this skill difficult and have not achieved as much as I would like in this area
D. I have not achieved this skill.

Question 39

Choose the statement which best reflects your attitude to the group work in this subject:

A. The groups worked well and I learnt a great deal

B. I prefer not to work in groups, but feel the experience was nevertheless a positive one
C. I prefer not to work in groups, and despite trying feel that they failed to help us learn as some people do little or no work
D. I prefer not to work in groups, and feel that it did not help us to learn anything.

Question 40
What mark did your group achieve for your presentation (/15)?

Question 41
Do you have any further comments which you feel would promote a more instructive learning environment in this subject?

Question 42
If you wish to be interviewed about your learning experiences in this subject, please enter your name and contact details below.

Thank you for your help in evaluating this subject. Your efforts are greatly appreciated. Good luck in the exam.

Signed:
(Lecturers names supplied)

Appendix 7.4: Spring 2005 Final Exam and teacher/marker comments

Please see print copy for Appendix 7.4